JNCC Report 742

Earth observation for habitat mapping and assessment in Scotland: status and opportunities in the context of 30x30

Pauline Burke, Paul Robinson, Lisa Hecker,
Christian Christodoulou-Davies and Philippa Vigano

August 2023

ISSN 0963 8091

JNCC's report series serves as a record of the work undertaken or commissioned by JNCC. The series also helps us to share, and promote the use of, our work and to develop future collaborations.

For further information on JNCC's report series please contact:
Joint Nature Conservation Committee
Quay House
2 East Station Road
Fletton Quays
Peterborough PE2 8YY
https://jncc.gov.uk/

Communications@jncc.gov.uk

This report should be cited as:
Burke, P.[1], Robinson, P.[2], Hecker, L.[2], Christodoulou-Davies, C.[3] & Vigano, P.[4]. 2023. Earth observation for habitat mapping and assessment in Scotland – status and opportunities in the context of 30x30. *JNCC Report No.742*, JNCC, Peterborough, ISSN 0963-8091
https://hub.jncc.gov.uk/assets/cde8edea-3cd2-4229-8d75-e36689b9935f

Author affiliation:

[1] JNCC, Inverdee House, Baxter Street, Aberdeen. AB11 9QA.

[2] JNCC, Quay House, 2 East Station Road, Fletton Quays, Peterborough. PE2 8YY.

[3] NatureScot, Great Glen House, Inverness. IV3 8NW.

[4] NatureScot, Battleby, Redgorton, Perth. PH1 3EW.

Acknowledgments:
This project was funded by Scottish Government through the Simple ARD Service. We would like to thank colleagues at JNCC and NatureScot for their input to this work, including Gwawr Jones (JNCC) for undertaking quality assurance of the report.

Evidence Quality Assurance:
This report is compliant with JNCC's Evidence Quality Assurance Policy
https://jncc.gov.uk/about-jncc/corporate-information/evidence-quality-assurance/

The views and recommendations presented in this report do not necessarily reflect the views and policies of JNCC.

Summary

This report reviews the status of earth observation (EO) for habitat mapping and assessment in Scotland in the context of 30x30. The 30x30 initiative pledges that by 2030, at least 30% of our land and seas should be protected or conserved for nature. In Scotland, and indeed more widely, there is increasing focus around how 30x30 will be achieved alongside other sustainable development and environmental priorities, and how nature recovery within and around protected or conserved areas should be monitored and assessed.

Habitats are important indicators of biodiversity, and understanding their extent, condition, and how they change over time, is likely to play an important role in delivering 30x30. In recent years there has been a shift towards greater use of EO data for habitat mapping. This is reflected for instance through the Scotland-wide land cover map, with similar initiatives across other UK nations. However, there is still a need for clarity around the current status of EO for habitat mapping – what levels of detail are feasible through EO, and what are the most appropriate methods for EO mapping? Other questions exist around quantifying habitat change and assessing condition, both of which are key needs in measuring progress towards the 2030 target. Assessing habitats in such ways are one means by which the effectiveness of 30x30 sites can be measured. This report addresses these issues by presenting a review of the current status of EO, considering EO data sources and classification approaches; data resolution and scale; spatial frameworks; computational and software resources; and uncertainty and the requirements for field reference data. Habitat change and condition assessment are also reviewed. Context is provided around habitat classification systems, including EUNIS as Scotland's preferred classification system for terrestrial habitats, as well as those used elsewhere across the UK.

A number of recommendations are highlighted. This includes the initial need to identify requirements and objectives relating to 30x30, so monitoring strategies can be appropriately designed to meet these needs. In turn, this will inform the scope, scale and data requirements around EO-based assessment of habitat for protected or conserved areas. The review identifies the predominance of machine learning EO classification approaches. There is also growing use of radar (SAR) and LiDAR datasets in habitat classification. These can provide additional information on vegetation structure, and in the case of SAR, can offer higher frequency data collection, which may compensate for the scarcity of cloud-free optical imagery, which is especially problematic in Scotland. There is also growing interest around very high resolution (VHR) satellite imagery which offers potential for more detailed habitat mapping. The need for high quality reference data to train models and validate results is a crucial element. This can be costly and challenging to achieve but should be considered at an early stage in the design of any habitat mapping or change assessment project.

There are also ongoing developments around quantifying habitat change, particularly concerning national-scale habitat mapping initiatives and quantifying year-on-year change. However, assessing habitat condition is currently more challenging. While the literature around assessing habitat condition is growing, most approaches are specific to particular habitats or sites and have not been more widely applied. However, both habitat change, and condition are likely to become increasingly important in terms of measuring the effectiveness of 30x30 actions.

In all cases, to identify the most suitable data and methods, it is crucial to start by considering what habitats, or characteristics, are of importance for assessment and monitoring. While EO datasets can be very large, and require advanced computational resources to process, manage and analyse, the availability of cloud computing solutions and open-source software presents good opportunities to leverage greater value.

Contents

1 Introduction

1.1 Habitat mapping

Habitats are important indicators of biodiversity, especially at landscape scale (Bunce *et al.* 2013), and are essential for informing management options in nature reserves and other protected or conserved sites (Nagendra *et al.* 2013). Habitats are closely linked to species and species groups sharing the same ecological requirements and can provide a framework for assessing species behaviour and wellbeing (Bunce *et al.* 2013). Understanding the extent of particular habitats and being able to reliably monitor changes in condition is important for conservation and management purposes, as well as in the wider context of reversing biodiversity decline. This is reflected through the legal and policy frameworks. For example at European level, habitats identified through Annex I of the Habitats Directive form the basis for sites protected through Special Areas of Conservation (SACs), forming part of the Natura 2000 network alongside Special Protected Areas (SPAs) designated under the Birds Directive. UK Natura 2000 sites have now been transposed to the UK National Site Network, and also form part of the Bern Convention's Emerald site network. Additionally, wetland habitats are protected under the Ramsar Convention. At the highest political level, the Convention on Biological Diversity (CBD) recently agreed the Kunming-Montreal Global Biodiversity Framework at COP15, which includes specific targets and national obligations relating to the protection of species and habitats. This includes the '30x30' initiative, which pledges that by 2030, at least 30% of our land and seas should be managed as protected or conserved areas.

1.2 30x30

The 30x30 initiative was first proposed in 2019 (Dinerstein *et al.* 2019), and launched by the High Ambition Coalition (HAC) for People and Nature in 2020, with more than 100 countries, including the UK, having committed to the initiative by October 2022. With subsequent inclusion in the CBD Kunming-Montreal Framework, 30x30 has become a major focus in biodiversity conservation at global, regional and country levels. In Scotland, 18% of land (including freshwater and coastal areas) is protected through designated sites including SACs, SPAs, SSSIs, National Nature Reserves and Ramsar sites (increasing to 23% when National Parks are included), with 37% of Scotland's seas protected (NatureScot 2023). In Scotland, the UK and more widely, there is focus around how the 30x30 target should be implemented, considering which areas should be prioritised and how this should be undertaken; how this can be achieved alongside other sustainable development and environmental priorities; and how nature recovery within and around protected or conserved areas can be monitored and assessed. These are challenging considerations which must also evaluate other competing requirements relating to land management and climate adaptation strategies, such as net zero targets.

Mapping and monitoring habitat extent and change is central to addressing many of the questions around effective management and reporting of protected or conserved areas (Jongman *et al.* 2019). The identification of potential sites for inclusion within 30x30 requires robust evidence that the areas chosen are of particular importance for biodiversity. Whether or not a site meets the criteria for a protected or conserved area will in many cases be informed by the presence of certain habitat types or a complex of habitats that may not be possible to identify from existing habitat maps. The increased ability to identify habitats, by a comparatively low-cost means, at a landscape or national scale is required not just for individual site identification but also to provide information on how well a site is connected to its broader landscape and neighbouring sites. The importance of 30x30 sites being well connected, and so needing to look outwith the boundaries of protected or conserved areas, is as important as reaching the goal of at least 30% protected or conserved.

Monitoring and reporting of the suite of sites is another key requirement for delivery of 30x30 and is likely to be an area where earth observation (EO) based approaches can play a significant role. Mapping and monitoring habitat extent and change on protected and conserved areas will be a key response metric that allows for a greater understanding of whether active management has led to desired outcomes. As the latency of reporting change from EO-derived data reduces, then it can increasingly be used to instruct management when change is identified.

The journey to 30x30 is a transformative one where protected and conserved sites will have increased complexity and flexibility, ensuring they have the tools to adapt to an uncertain and changing climate. The suite of sites contributing must still represent the breadth of Scotland's habitats and it will be important to regularly report on the comprehensiveness of the 30x30 suite. EO data is likely to play a significant role in the ability to report on habitats and their extent and condition within the 30x30 sites.

1.3 Earth observation for habitat mapping

Airborne- and satellite-based EO approaches have been extensively explored over several decades for mapping and monitoring of the natural environment. Over the last decade, momentum around EO for habitat mapping has gathered pace, underpinned by the increasing availability of high and very high resolution (VHR) satellite imagery, and developments in machine learning and computational resources. These aspects are discussed further in Section 3.

Habitats are crucial elements in the quantitative assessment of biodiversity across a range of scales (Agrillo *et al.* 2021). Refining this further, Mikula *et al.* (2021) identify three main areas where EO can contribute towards habitat assessment:

- Habitat distribution.
- Habitat change detection.
- Habitat condition.

Available methods to assess these three areas are variable in terms of our understanding and readiness to apply them, but reliable, up-to-date habitat maps, at appropriate scales are fundamental to enabling such assessment. While these observations present a broad context for how EO can contribute, identifying the underlying policy objectives is imperative in focussing expectations and scope. This allows the identification of specific priorities, constraints, and expected deliverables. In the context of EO, this will usually come down to available resources (monetary, skills, personnel, computational) and a trade-off against the achievable scale and temporal frequency of mapping. For example, the requirements and resources related to producing a national-scale habitat map will inevitably lead to constraints around scale or level of detail. In contrast, habitat mapping of a single site or even a network of protected or conserved areas may allow for more detailed mapping and more bespoke analysis related to the site characteristics. Another important consideration relates to the acceptable accuracy and uncertainty of habitat maps or change detection products. This is discussed further in Section 5. In many cases it is desirable to be able to monitor how a landscape is changing over time, and how this may relate to habitat condition or quality, and the impact of land management practices. This is likely of high importance in the context of protected or conserved areas and 30x30, where it may be desirable to assess not only changes within these areas but also in the adjacent landscape. These aspects are further considered in Section 6.

Together these elements highlight the need for careful design and planning around the use of EO in mapping and monitoring of habitats and the wider landscape in and around protected or conserved areas.

1.4 Habitat mapping for Scotland

An early map of what we would now describe as land cover was produced by Dudley Stamp in the 1930's and is called the Land Utilisation Survey (Stamp 1931). This was achieved by teams of people, largely students, visiting all areas across Great Britain and manually colouring copies of Ordnance Survey 1-inch to 1-mile maps. The data has been digitised and could be a useful resource for long term change assessment at a broad scale.

More detailed data was generated in the Land Cover Scotland 1988 dataset which utilised manual interpretation of aerial photography to generate outputs which are still available and regularly used today, as they are considered of particularly high-quality.

Further Land Cover Maps (LCM) have been produced by UKCEH using automated analyses of satellite data (discussed further in Section 4). These products have evolved in terms of their production and format of the output, since the first publication in 1990. This evolution is unsurprising given the huge change in available data, analytical techniques, and computing capabilities over that period of time. The LCM are updated on an annual basis and UKCEH have started to generate test datasets describing change over time between certain iterations of the datasets.

Several iterations of the broad scale CORINE land cover map have been produced by the EU with later versions coming under the banner of the Copernicus Programme. The first three iterations, from 1990, 2000 and 2006 were created for the UK by converting the UKCEH LCM into the CORINE specifications, but the last two versions (2012, 2018) have been generated for the UK by following the main CORINE production process. They are intended to be broad scale, with a minimum mapping unit (MMU) of 25 ha and are unsuitable for detailed mapping purposes. A separate change layer with MMU of 5 ha was produced for the UK for the period 2012-2018 (Cole *et al.* 2021).

Most recently, a land cover map (known as the 'SLAM-MAP') has been generated on a national scale for Scotland, and is discussed further in Section 4.1.

A number of ground based habitat surveys were brought together as the Habitat Map of Scotland (HabMoS). These data, collected over many years for different end uses were all converted to EUNIS and the European Directive Annex I habitats. This has the benefit of being based on detailed, specifically designed habitat surveys but does not have complete geographical coverage and is based on data collected over a range of time periods.

All these datasets have been produced following different methodologies and have varying aims and resolutions showing how important it is for users to properly understand these datasets and associated uncertainties, to ensure they are used appropriately.

1.5 Scope and methodology

This review sets out to assess the current status of EO-based habitat mapping, including habitat change and condition assessment. The principal focus is around reviewing current EO data sources and analytical approaches and considering these in the context of needs for the developing 30x30 policy agenda in Scotland.

The work was undertaken through literature review. This was initiated by reviewing a number of published studies and reports highlighted by NatureScot as well as JNCC colleagues. A semi-structured literature review was then conducted on the Scopus academic database (undertaken late October 2022). A number of keywords were identified (e.g. Earth Observation; Remote Sensing; Sentinel-2; VHR; Machine Learning, Habitat Mapping) and used in combination to identify potentially relevant articles from peer reviewed journals,

selecting only those published in or after 2010. This was further narrowed down by rapid review of the abstracts to eliminate those of less relevancy. This produced a total of 104 articles, some of which proved unavailable as they were not open access. Available articles were further prioritised as being of high, medium or low priority. This process led to around 30 articles being reviewed, with a small number of additional articles (including older, key literature) subsequently included to support various aspects of the review as it developed.

2 Habitat classification systems

2.1 Land cover and habitat classification

The capture, analysis and interpretation of earth observation and ecological data requires consistent definitions of features on the ground (Brownett & Mills 2017). Habitat classification systems are often closely linked to land cover classification, although it is important to distinguish between the two. Land cover classification tends to represent the characteristics of the Earth's surface at relatively broad scales. Habitat classification can be considered more detailed, placing emphasis on identifying vegetation units, defined by their species composition (Mikula *et al.* 2021). Until recently, the majority of EO studies focussed on land cover, as opposed to habitat mapping which is more challenging to undertake (Nagendra *et al.* 2013).

Examples of land cover classification systems include the FAO Land Cover Classification System (LCCS) at global scale, and CORINE Land Cover (CLC) for Europe at regional scale. Examples of habitat classification systems include General Habitat Categories (Bunce *et al.* 2008) at global scale, the European Nature Information System (EUNIS) developed by the European Environment Agency (EEA), and the Phase 1 Habitat Classification System in the UK. From this basis, land cover and habitat maps have been developed across a range of scales. Some of these are based on the formalised classification systems mentioned above, while others are more bespoke or generalised classification systems related to the application.

Traditionally, habitat maps are produced through surveys undertaken by field ecologists, often in combination with aerial photography interpretation (API) to allow mapping across larger extents. While such approaches continue to play an important role, the widespread availability of EO data has seen the development of alternative approaches. Following the advent of medium resolution multispectral satellite sensors offering global coverage (e.g. Landsat, SPOT), it became increasingly possible to generate EO-based maps, with opportunities for regular updates and change monitoring. Owers *et al.* (2021) noted the importance of such maps in supporting sustainable development activities, and their relevancy in establishing baseline conditions for monitoring change across a range of scales.

More recently, the Copernicus programme is delivering freely available high resolution satellite imagery, further augmented by the availability of commercial VHR satellite data, which together offer increasing potential for producing detailed habitat maps (Tomaselli *et al.* 2013). These developments have led to efforts to translate from commonly used national, regional, and global land cover classification systems to habitat classification systems, which offer greater relevancy for biodiversity monitoring (Jongman *et al.* 2019). Some implementations use land cover classification frameworks as a basis for habitat mapping (e.g. Planque *et al.* 2020). However, translating between land cover and habitat classification can be challenging due to differences in definitions and criteria, as discussed in detail by Tomaselli *et al.* (2013). Furthermore, most land cover and habitat classification schemas are designed for field mapping or visual image interpretation and are therefore not optimised for mapping using EO approaches (Morton & Rowland 2015).

2.2 Habitat classification systems in the UK

The UK has a complex legacy of habitat classification systems, all of which have been designed for use in field surveys. As they have not been designed with EO applications in mind (with the exception of UKHab), there generally must be some form of compromise made when being applied in that context.

The Phase 1 habitat classification system was developed to provide a standardised system for recording semi-natural vegetation and other wildlife habitats through field survey. It was first published in 1990 after many years of development. The Phase 1 approach has been extensively used for local ecological assessment, as well as country-wide surveys, and is widely used for environmental impact assessments. The system is comprised of ten broad categories, with more detailed sub-categories, identifying a total of 155 habitats.

The National Vegetation Classification (NVC) was published between 1991 and 2000 and aimed to produce a comprehensive and highly detailed classification of the plant communities and vegetation of Britain (thus excluding Northern Ireland). The NVC has formed a basis for detailed ecological site surveys and assessments undertaken across Britain by a range of organisations, including the Country Nature Conservation Bodies. It also formed the basis for the interpretation of Annex I habitats under the EU Habitats Directive.

The UK Biodiversity Action Plan, published in 1994, identified a range of semi-natural habitats, termed UK BAP Priority Habitats, which were identified as being at greatest risk and requiring conservation action. The original list was produced between 1995 and 1999 and revised in 2007. Additionally, UK BAP Broad Habitat types are defined, with each class composed of one or more of the Priority Habitats. The Broad Habitats are intended to allow for evaluating habitats in whole-UK context. The UK BAP Priority Habitats have remained important, although the UK BAP no longer has legislative stature, having been succeeded by the UK Post-2010 Biodiversity Framework and related country-level biodiversity strategies.

The UK Habitat Classification (UKHab) was published in 2018 and provides a detailed hierarchical classification system for the UK. UKHab is designed to provide a rapid system for classifying habitats, which is applicable for both field surveys and EO approaches. It builds on existing classification schemes used across the UK, enabling interoperability between these different approaches, and includes translation schemes for Phase 1, NVC, UKBAP and EUNIS. It is comprised of primary habitats with an underlying hierarchy of five levels. This allows for broad scale surveys, as well as detailed habitat mapping. Merrington *et al.* (2021) observe that UKHab offers an opportunity for EO in that habitat maps can be devised for a specific level of detail based on related land management applications. Currently however, usage is limited by licensing constraints.

At regional level, EUNIS was developed by the EEA to provide a consistent system for pan-European classification of marine and terrestrial habitats. It is a comprehensive, hierarchical classification scheme, with criteria for identification of habitats at the first three levels (Moss 2008). It is distinct from Annex I habitats of the EU Habitats Directive but identifies correspondences to these. It allows for identification of habitats at very broad scale (e.g. Level 1), as well as detailed scale (e.g. Level 3 and beyond). NatureScot use EUNIS as the standard terrestrial habitat classification system in Scotland.

3 Earth observation for habitat mapping

3.1 Habitat mapping approaches

Habitat mapping has conventionally been undertaken by field ecologists, often in combination with API, which can be considered an EO technique in its own right. While field surveys are founded on expert knowledge, they are labour intensive, costly, and unable to cover large extents (Merrington *et al.* 2021). There are further disadvantages in that vegetation and habitat mapping in the field can be relatively subjective and has been found to differ significantly between individual surveyors (Hearn *et al.* 2011), particularly in the case of mosaic habitats and complex upland environments (Lucas *et al.* 2007). This can lead to inconsistency and increased levels of uncertainty. API approaches can reduce the field burden and enable mapping over larger areas. API makes use of vertical aerial photography, captured through photogrammetric approaches to enable stereo viewing, which aids interpretation. However, as discussed by Lucas *et al.* (2007) aerial image capture is extremely expensive, and due to limited suitable weather conditions, especially in countries such as the UK, it is challenging to ensure that imagery is captured at an optimum time of year for discrimination of key habitats. API also relies on expert ecological interpretation, and in a similar vein to field survey, mapping may be subjective, leading to relatively high uncertainty in comparison to other EO approaches (e.g. Pesaresi *et al.* 2022). While attention has turned increasingly to the use of satellite imagery for habitat mapping, conventional approaches are by no means obsolete. Indeed, field survey has an important role to play in terms of validating contemporary EO methods, as discussed in Section 5.

EO-based habitat mapping requires careful consideration at an early stage around the design and implementation of a suitable approach. Fundamentally the goal is to use EO-derived data to identify and assign habitat classes, or related information (e.g. habitat change, habitat condition) to produce reliable and accurate maps or related metrics. The scope and objectives of the habitat mapping or monitoring (e.g. level of detail and MMU, spatial extents to be covered, frequency of updates for change monitoring, etc.) are crucial for determining the most appropriate EO-based methodology. This relates to considering the suitability of different types of EO data, including aspects such as availability and cost. This also relates to considerations around data resolution, the integration of other relevant datasets, and the most appropriate classification method, as well as how the classification will be implemented in a spatial context (spatial framework). Software and computational resources are also a key consideration. These various elements are discussed in the following sections.

3.2 Classification of EO data

3.2.1 Pixel-based classification

Since the early days of remote sensing from satellites, pixel-based classification has been the primary means of assigning meaning and extracting information from satellite imagery. This includes land cover and habitat mapping, where statistical classifiers such as maximum likelihood classification have been the established approach (Brownett & Mills 2017). In a supervised classification strategy, this relies on the user identifying (labelling) sites or locations within an image which correspond to a pre-defined land cover or habitat class. The classifier then analyses the spectral distribution of these samples and identifies other pixels in the image with similar characteristics to assign classes across the whole image. Unsupervised pixel-based approaches examine the spectral characteristics of the image, and then attempt to classify pixels as belonging to discrete classes, which the user is then able to label. In the scope of land cover and habitat mapping, pixel-based classification has

largely been superseded by object-based methods, where the characteristics of individual pixels are aggregated into objects as relating to real-world entities on the ground.

3.2.2 Object-based classification

It has long been recognised that features in the real-world bear little similarity to pixels, and it is intuitive that ecological features can benefit from consideration of spatial patterns where it is likely that neighbouring pixels will relate to the same feature (Blaschke & Strobl 2001). This has led to geospatial object-based image analysis (GEOBIA), more commonly termed OBIA, becoming mainstream in classification of satellite imagery for environmental monitoring. The increasing availability of high-resolution satellite imagery has accelerated this, as pixel size is often significantly smaller than the features under consideration, which supports analysing groups of pixels as objects (Blaschke 2010), illustrated in Figure 1.

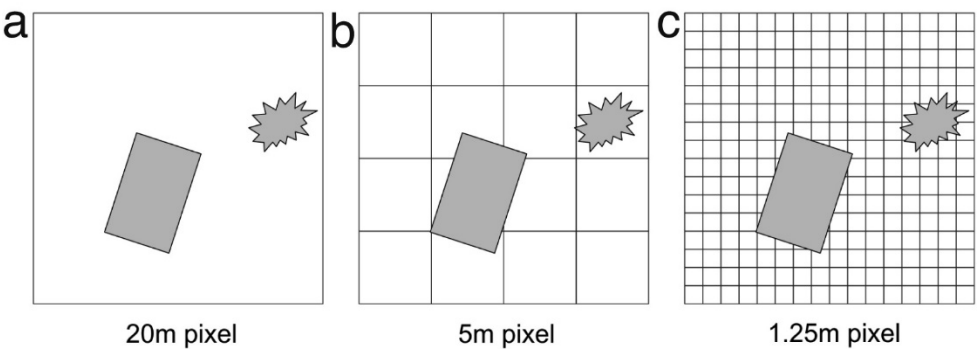

Figure 1. Relationship between pixel and feature size at different spatial resolutions (Blaschke 2010).

OBIA attempts to identify meaningful real-world features in imagery by examining spectral information alongside characteristics relating to shape, size, and texture (Hay & Castilla 2008; Smith & Morton 2010). This is implemented through image segmentation, where the image is divided into contiguous objects or segments relating to landscape features, such as habitat patches, fields, woodland, etc. These can then be classified through a range of supervised or unsupervised approaches which relate the characteristics of the segmented objects to, for example, habitat classes. OBIA offers the added advantage of multi-scale analysis, where image segmentation can be adaptable to the scale of analysis or features under consideration (Blaschke 2010). This is particularly advantageous for habitat mapping from satellite imagery, where adaptable scale hierarchies can support combined identification of larger homogeneous habitats and more detailed fine scale habitats or habitat mosaics, as typical of the UK landscape (Sideris *et al.* 2020).

3.2.3 Machine learning classification

Over the last decade, machine learning approaches have seen increased uptake for land cover and habitat mapping. This reflects a transition from more conventional statistical or parametric approaches to machine learning methods, although Maxwell *et al.* (2018) observed a need for greater training within the user community as to how best to utilise and implement these approaches. The advantages of machine learning methods include their capacity to model complex class signatures, and ability to accept additional predictor inputs (e.g. ancillary datasets as discussed in Section 3.8) (Maxwell *et al.* 2018).

The random forests (RF) method (Breiman 2001) is particularly popular for EO applications. It has generally been shown to outperform other methods in terms of classification accuracy (Belgiu & Drăguţ 2016), and has been applied in a wide range of vegetation and habitat mapping studies (e.g. Agrillo *et al*. 2021; Andries *et al*. 2021; Pesaresi *et al*. 2022). RF is a supervised classifier which employs a ‘forest’ of decision trees, based on the concept that

the collective decision of an ensemble of weak predictors presents a strong solution (Morton & Rowland 2015). RF also offers the benefit that it enables the identification and ranking of variables in terms of their performance in discriminating between classes (Belgiu & Drăguţ 2016). This is attractive given the complexity and high dimensionality of EO data. Other commonly utilised machine learning classifiers include support vector machines (SVM), single and boosted decision trees, artificial neural networks (ANN), and k-nearest neighbours (k-NN). Maxwell *et al.* (2018) presented a comprehensive review of machine learning methods for EO which addresses many of the key considerations around selecting an appropriate classifier.

3.3 Considerations around resolution and scale

3.3.1 Spatial resolution

Whether pixel- or object-based classification methods are used, the spatial resolution of EO data needs to be considered carefully in comparison to the features or objects to be mapped. In the case of pixel-based approaches, the pixel size is the primary constraint in terms of the scale of features which can be resolved. However, the traditional concept of a MMU is important to constrain the smallest feature on the ground which should be identifiable in the mapping, especially in the case of OBIA. For example, in the Living England maps, a MMU of 3 pixels (300 m^2) is defined, and objects smaller than this are merged with a neighbouring object (Kilcoyne *et al.* 2022). In the case of the UKCEH Land Cover Maps, land parcel objects have a MMU of 0.5 ha (5,000 m^2), although the maps are produced and made available through a pixel-based classification with a spatial resolution of 10 m^2 thus preserving a greater level of detail. In the case of other EO datasets, such as LiDAR, spatial resolution relates to the point density, which may equate in turn to pixel size if the LiDAR data is converted to a raster digital elevation model (DEM), digital surface model (DSM) or canopy height model (CHM).

3.3.2 Spectral resolution

Considering that satellite imagery is often the primary EO data source for habitat mapping, resolution relates not only to the spatial aspect, but also spectral. The most widespread sensors currently used for land cover and habitat mapping are multispectral, which typically offer 3 to 10 spectral bands covering the visible, near infrared and into the short-wave infrared parts of the electromagnetic (EM) spectrum. Figure 2 shows a comparison of the spectral bands for Landsat 7, Landsat 8, and Sentinel-2.

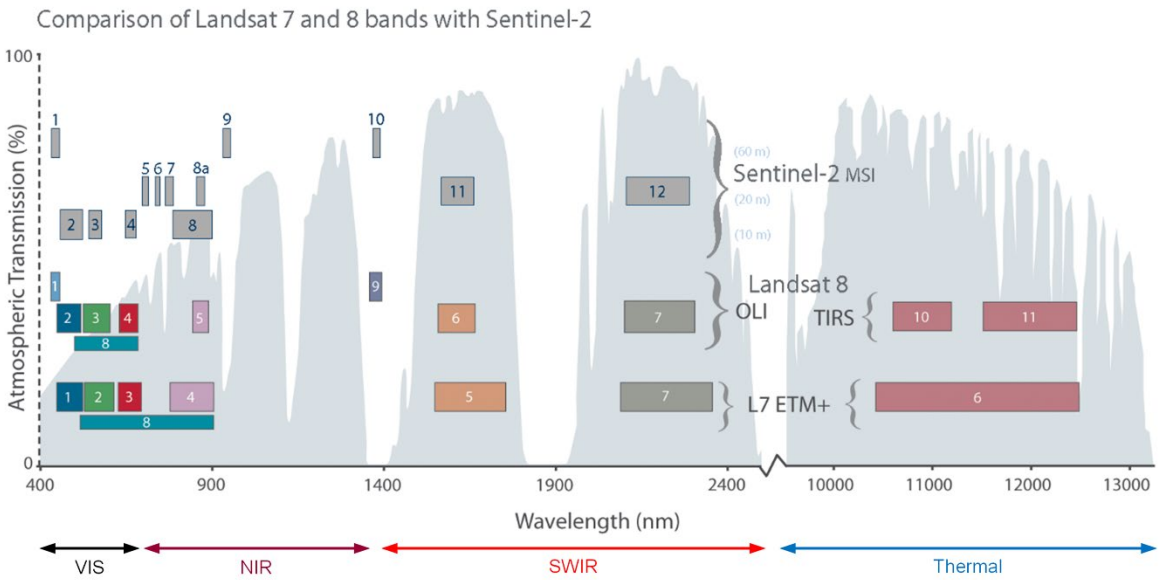

Figure 2. Comparison of Landsat 7, 8 and Sentinel-2 spectral bands. Adapted from https://landsat.gsfc.nasa.gov/wp-content/uploads/2015/06/Landsat.v.Sentinel-2.png.

The near-infrared portion of the spectrum is known to be especially useful in discriminating vegetation characteristics, including seasonal traits and responses to stress (e.g. drought, disease) which are often captured by deriving vegetation indices (see Section 3.4). Most multispectral and hyperspectral sensors include specific bands in this portion of the spectrum due to their utility for many terrestrial monitoring applications. The applicability of different sensors for habitat mapping is discussed in more detail in Section 3.7.

3.3.3 Radiometric resolution

Radiometric resolution, referring to the 'bit depth' of the imagery is important to consider as this dictates the information content at pixel level (Nagendra *et al.* 2013). This means that data captured at higher bit depths is better able to resolve finer spectral detail. 12-bit radiometric resolution is commonplace (e.g. Landsat 8 and PlanetScope) and Sentinel-2 is delivered at 16 bit resolution.

3.3.4 Temporal resolution

Temporal resolution is an important consideration and relates to the frequency at which EO sensors can collect data over the same area. Satellite sensors are continuously orbiting the Earth and are typically able to collect data over the same location with a repeat frequency of several days or weeks. This contrasts with surveys conducted by methods such as LiDAR and aerial imagery, which are specifically commissioned for acquisition over defined areas, and would at best be able to offer repeat capture over intervals of several months or more likely years. These aspects will be discussed further in Section 3.7 relating to specific sensors. Temporal resolution is particularly important in the context of assessing habitat condition and change over time (Section 6).

3.4 EO indices

Indices derived from EO data are widely used in environmental monitoring for applications including agriculture, forestry, geology, soil science and vegetation analysis. Vegetation indices can be useful in characterising vegetation health and phenology. Spectral indices can be considered as products, derived through a calculation or ratio applied to two or more spectral bands. This calculation is applied to each pixel in the image to derive the related

index value. Some of the more commonly applied indices used in land cover and habitat mapping are (USGS 2023):

- Normalised Difference Vegetation Index (NDVI)
    - A measure of vegetation 'greenness' useful for identifying healthy vegetation
- Enhanced Vegetation Index (EVI)
    - Similar to NDVI, but corrects for some atmospheric influences and canopy background noise in areas of dense vegetation
- Normalised Difference Moisture Index (NDMI)
    - A measure of vegetation moisture content
- Soil Adjusted Vegetation Index (SAVI)
    - An alternative to NDVI, correcting for influence of soil brightness in regions of low vegetation cover

Additionally, radar indices are also increasingly being used for vegetation analysis – for example the Radar Vegetation Index (RVI). However, the interaction of the microwave signal with vegetation and other contributory factors, such as soil scattering, is complex and radar indices such as RVI are less well-understood in comparison to indices derived from optical imagery (Szigarski *et al.* 2018).

As well as providing insights on vegetation characteristics directly, the application of machine learning classifiers such as RF means that indices are increasingly being used in the classification process as predictor variables (e.g. Agrillo *et al.* 2021; Pesaresi *et al.* 2022). Index values can also be used directly to detect land cover or habitat change, by evaluating changes to their values over time. For example, JNCC have developed the Landscape Evaluation Tool (Lightfoot *et al.* 2021), a web-delivered app which allows assessment of changes to indices derived from timeseries of Sentinel-1 and Sentinel-2 imagery, including NDVI and NDMI. This is designed to enable assessment of land parcel changes against baseline habitat maps through expert interpretation by land managers and site specialists and was recently evaluated by NatureScot for monitoring applications within protected sites (Black *et al.* 2023).

3.5 Spatial frameworks for habitat mapping

A spatial framework is a set of land parcels or segmented objects used as the basis for assigning land cover or habitat classes derived from EO sources (Morton & Rowland 2015). Thus, a spatial framework provides the spatial structure for applying a habitat classification. This is inherently related to the scale or level of detail at which meaningful classes will be assigned (Blaschke 2010). Such a framework can be derived through image segmentation, for example as part of an OBIA approach (e.g. Living England) or may be generated form digital cartography (e.g. UKCEH Land Cover Maps). Defining an MMU, as discussed in Section 3.3.1, is an important early step in constraining the size of land parcel objects to be used within the spatial framework. The challenges around developing suitable spatial frameworks for habitat mapping, especially in the context of change detection, are discussed at length by Barber and Robinson (in press). In brief though, two key issues need to be faced. Firstly, how frequently to update that spatial structure to represent changes in the geographical objects themselves, particularly changes that relate to parts of individual polygons. Secondly, once a spatial framework has been updated, decisions need to be made in defining how that change can be analysed between the different iterations of the map. Particularly considering the size of change that can be reliably identified using the available data and how the analysis will handle smaller areas.

3.6 Software and computing infrastructure

3.6.1 Overview

The exploitation of high spatial and temporal resolution EO data has not only been advanced through use of OBIA and machine learning, but critically is underpinned by access to advanced computational resources. Over recent years, developments such as cloud computing and access to high performance computing infrastructure have transformed the way in which EO analyses can be undertaken. Conventionally, the classification and analysis of EO data relied on expert knowledge and expensive, proprietary software. Over the last decade, there has been a huge shift towards open-source software, making EO data more widespread and accessible. This has coincided with the proliferation of cloud computing platforms, such as Google Earth Engine, offering readily available analytical resources which can be implemented in relatively 'black box' fashion by application specialists rather than remote sensing scientists. Machine learning algorithms, previously the exclusive domain of computing scientists, are now available for use 'off-the-shelf'. While these developments build momentum around leveraging the potential of EO, they must be treated with caution. It is important that informed decisions are made around the use of appropriate classification models, as well as understanding requirements for training data and appropriate validation approaches.

3.6.2 Software

Over the last decade or more, open-source software has become widespread, bringing about significant cost savings over expensive licensed software, and increasing the accessibility of EO data. Open-source software can offer greater flexibility for developing bespoke user workflows for EO data processing and is often underpinned by an active user community. However, in some cases open-source software may be less user-friendly compared to commercial alternatives and technical support may be limited.

While many OBIA applications continue to make use of commercial software, especially Trimble eCognition, open source options, such as Orfeo ToolBox (OTB) also offers OBIA. Sideris *et al.* (2020) compared eCognition, OTB and GRASS GIS for segmentation of Sentinel-2 imagery for an area of varied land cover types in the Scottish borders for which ground surveyed land cover parcel polygons were available. All three software tools provided acceptable segmentations, although eCognition was found to better facilitate multi-scale segmentation and processing over larger areas. OTB and GRASS were well-suited to detailed segmentation over more limited extents, with the benefit of significantly lower costs through the open-source nature of these tools (Sideris *et al.* 2020). A comparison of segmentation methods from the open source scikit-learn package is presented by (Merrington *et al.* 2021). An important consideration is the trade-off between level of detail and practicality. This relates to the scale or size of the segments, the spatial extent over which the segmentation is performed (e.g. local, regional, national), and the cost in terms of time and manual effort needed to refine or edit a segmentation. There is typically a processing bottleneck around the segmentation process, including the calculation of statistics for each segment, which has implications for OBIA implemented across large spatial extents (Merrington *et al.* 2021). Pixel-based classification approaches assign class values at the pixel level, and therefore avoid the need for image segmentation and associated processing demands.

Considering the trend towards machine learning classification of EO data for habitat mapping, most reported examples implement open source toolsets, including various *R* implementations (Agrillo *et al.* 2021; Kilcoyne *et al.* 2022; Pesaresi *et al.* 2022; Schmidt *et al.* 2018), Weka (Morton & Rowland 2015), and scikit-learn (Merrington *et al.* 2021). QGIS also

offers implementations of several commonly used machine learning classifiers. Table 1 highlights the software availability of machine learning classifiers (Maxwell *et al.* 2018).

Table 1. Software implementations of machine learning methods including k-nearest neighbours (*k*-NN), artificial neural networks (ANN), decision trees (DTs), random forest (RF) and support vector machines (SVM) (Maxwell *et al.* 2018**)**. An X indicates algorithm used by software package.

Software type	**Software name**	**Algorithm**					
		k-NN	**ANNs**	**DTs***	**Boosted DTs**	**RF**	**SVM**
Geospatial software	ArcGIS 10.5/ArcPro 2	-	-	-	-	X	X
	eCognition Developer 9	X	-	X	-	X	X
	ENVI 5.4	-	X	-	-	-	X
	Erdas Imagine 2016	-	-	-	-	-	-
	TerrSet/IDRISI 18.3	X	X	X	-	-	-
	QGIS 2.18	X	X	X	X	X	X
Other software	Matlab 9.3	X	X	X	X	X	X
	R 3.4	X	X	X	X	X	X
	Python	X	X	X	X	X	X
	scikit-learn 0.19.1	-	-	-	-	-	-
	Weka 3.9	X	X	X	X	X	X

***** We define DTs as algorithms that generate a decision tree using machine learning. Our definition excludes expert-system decision trees, where the trees are developed from expert knowledge. Both ENVI 5.4 and Imagine 2016 have environments for encoding expert systems.

ArcGIS is sold by ESRI (Redlands, CA, USA), eCognition by Trimble (Sunnyvale, CA, USA), ENVI by Harris Geospatial Solutions (Melbourne, FL, USA), and TerrSet/IDRISI by Clark Labs (Worcester, MA, USA). Matlab is sold by MathWorks (Natick, MA, USA). QGIS, R, scikit-learn, and Weka are open-source software. QGIS, R, and scikit-learn are collaborative efforts; Weka is produced by the Machine Learning Group at the University of Waikato, New Zealand.

3.6.3 Computing infrastructure

Recent habitat mapping work undertaken across the UK has adopted a range of computational implementations. Living Wales currently uses a supercomputing facility to produce land cover and habitat maps. From the user perspective, this is soon to be delivered through a data cube approach. Refer to Wilson *et al.* (2021) for an evaluation of the Open Data Cube. Living England uses Google Earth Engine to source, filter and prepare Sentinel imagery for subsequent use in the classification workflow, using an open-source *R* package to perform the RF classification. JNCC's Land Evaluation Tool communicates change at land parcel level through spectral indices (Lightfoot *et al.* 2021). This is based on the JASMIN supercomputing facility operated by the Science and Technology Facility Council, and also makes use of Amazon Web Services (AWS) cloud computing resources.

This landscape is expected to develop further in the near future, as the National Centre for Earth Observation (NCEO) has recently announced the development of an EO Data Hub with ambitions for this to provide a single UK portal for accessing EO data, and offering

embedded computer resources (utilising JASMIN and the CEDA archive). Furthermore, the Copernicus Data Space Ecosystem was launched in early 2023, providing re-vamped user access to Copernicus EO data, alongside online analytical resources based on cloud computing.

While these approaches may differ individually, the common theme is the use of high-performance cloud computing resources, where the EO processing is undertaken remotely from the user. This eliminates the need for local download, storage and processing of the huge data volumes associated with EO. The availability and uptake of cloud computing is another factor which has increased the accessibility of EO data for a broader user base.

3.7 EO data sources for habitat mapping

This section summarises the main types of EO data currently used in land cover and habitat mapping. While this summary is not exhaustive, and other sensors such as unmanned aerial vehicles (UAVs) have been investigated, these are the main sources which are likely to support habitat mapping over extended areas, with readily available data. Much of the work on habitat mapping implemented across the UK, and reported in the literature uses combinations of these data sources.

3.7.1 Aerial photography

Aerial photography can be considered as a form of VHR imagery. Section 2.1 discussed the use of API in conventional approaches for habitat mapping. However, aerial photography also been explored as dataset in its own right for EO-based habitat mapping. Merrington *et al.* (2021) compared Sentinel-2, WorldView-2, and aerial photography for upland habitat mapping in Scotland. They conclude that while 0.25 m aerial photography provided improved classification accuracies over 20 m Sentinel-2 data, multispectral 0.5 m WorldView-2 offered the highest accuracy, illustrating the value of spectral resolution over spatial resolution. In the Upland Mapping Pilot for Scotland an OBIA segmentation was applied to 0.25 m colour infrared aerial photography, and classification at EUNIS Level 3 or better was undertaken through a manual GIS-based mapping approach. While OBIA serves to accelerate the mapping process, the overall approach remains highly labour intensive. Another aspect concerning aerial photography is the need for radiometric calibration to deliver normalised digital number (DN) values to be used as ‘spectral’ information across mosaiced images. This was addressed in the Upland Mapping Pilot by Scobie *et al.* (2018). The retrospective nature of this correction is likely to introduce a level of uncertainty which could be avoided through use of VHR satellite imagery where more rigorous radiometric and atmospheric correction will be applied at product generation stage.

However, one of the main challenges relating to aerial photography is the cost and availability of data. As discussed in Section 3.3.4 aerial photography must be specifically commissioned through a commercial provider. There are high costs associated with mobilising a survey crew and aircraft. Furthermore, coordinating this to coincide with a suitable weather window (clear sky conditions) is extremely challenging, especially in northern and western regions of the UK (see also Section 3.1). Additionally, access to airspace can add further complications, particularly around busy commercial airports. These factors also make it extremely difficult and costly to achieve repeat surveys targeted at specific times of year, or indeed national coverage, which would be staggered across multiple time periods as dictated by suitable weather conditions.

In the UK, the Aerial Photography for Great Britain (APGB) programme provides orthorectified aerial imagery, including associated digital elevation models for public sector users. These data are available on a three-year update cycle, but as the final product is a

mosaic of imagery captured at different points in time (for the reasons explained above), it is not well-suited to habitat mapping applications over extensive areas.

3.7.2 Satellite imagery

High resolution multispectral imagery

Multispectral satellite imagery refers to sensors which collect data typically over 3 – 10 spectral bands (Section 3.3.2). These are relatively broad bands, which cover specific portions of the EM spectrum, typically from the visible to near infrared (e.g. Figure 2), although some sensors (e.g. Landsat) also incorporate bands in the thermal infrared region. Such sensors are passive and collect energy as reflected from the Earth's surface. In the context of environmental monitoring applications, high resolution multispectral sensors typically refer to those with a spatial resolution of less than 6 m – 30 m. Sensors in this category include Landsat 8 and Sentinel-2. These sensors typically offer global coverage, with relatively high re-visit frequency. For example, Sentinel-2 offers a combined re-visit frequency of two to three days over Europe. However, optical sensors are affected by cloud cover, which means that, the revisit period over much of the UK, and especially Scotland, is significantly reduced. This has implications for creating seasonal mosaics or undertaking time series analysis in support of habitat mapping and change assessment. While these sensors typically make data available free of charge, there are still costs associated with processing into ARD format for ready uptake by users.

Sentinel-2 offers multispectral imagery at 10 m spatial resolution (20 m for NIR and SWIR bands), and has been widely applied for land cover and habitat mapping (e.g. Agrillo *et al.* 2021; Mikula *et al.* 2021). Sentinel-2 forms the foundation for habitat mapping for Living Wales, Living England, Scotland (SLAM-MAP) and Northern Ireland (JNCC Living Maps approach). At 10 m spatial resolution, Sentinel-2 is best suited to habitat mapping at relatively broad scale – such as the UKBAP Broad Habitats classes used for Living England, or EUNIS Level 2 in the case of SLAM-MAP. However, with a swath width of 290 km and high temporal revisit frequency, it provides a good foundation for mapping over large extents at regional- and national-scales.

Many studies have examined Sentinel-2 against other EO data sources or have combined it with other datasets to enhance the classification process. As already mentioned, Merrington *et al*. (2021) compared Sentinel-2 to aerial photography and VHR imagery (WorldView-2) for mapping of broad habitats at UKHab Level 3 in an upland area in Scotland. They applied an OBIA segmentation and RF classification and report overall classification accuracies of 50% for Sentinel-2, 75% for aerial photography and 93% for WorldView-2. They observe that the lower spatial resolution of Sentinel-2 was likely not sufficient to resolve habitat classes at the object length-scales applied, although it should be noted that Sentinel-2 bands were down-sampled to a common resolution of 20 m which will have impacted this. Other approaches apply band-sharpening to bring all bands to a common 10 m resolution (e.g. JNCC's Simple ARD Service). While VHR may have potential to deliver higher classification accuracies compared to Sentinel-2, costs are also likely to have a major influence. VHR data is significantly more expensive than Sentinel-2, even factoring in costs required to process Sentinel-2 to ARD stage.

VHR imagery

Very high resolution (VHR) imagery typically refers to satellite sensors with a spatial resolution of 0.3 m – 6 m. In the civilian domain, these sensors are operated by commercial providers, and can normally be tasked to acquire on-demand coverage of specific regions. Users pay for access to data, normally at significant cost. Typically, VHR satellite imagery has tended to offer fewer spectral bands in comparison to other multispectral imagery, often

restricted to visible, with perhaps a single NIR band. Often a panchromatic band is acquired at higher spatial resolution and may be used to 'pan-sharpen' the multispectral image bands. While pan-sharpening can improve opportunities to undertake more detailed analysis using multispectral imagery, it must be treated with caution as there is also potential to introduce spatial and spectral distortions which may detrimentally impact subsequent analyses (Jones *et al.* 2020). More recent VHR sensors offer improved spectral resolution. For example, WorldView-2, offers eight spectral bands in the visible and NIR, with 0.5 m panchromatic (pan) spatial resolution, and 2 m multispectral resolution. Planet Labs operate one of the largest constellations of VHR satellites, with three distinct sensor groups: Planet Doves (most recent sensors offer 3 m, 8 bands), RapidEye (5 m, 5 bands) and SkySat (0.5 m, 5 bands). Other commonly-utilised VHR sensors include Pléiades (0.5 m pan, 2 m multispectral, 4 bands) and Vision-1 (0.9 m pan, 3.5 m multispectral, 4 bands). The latest SPOT satellites, SPOT 6/7 offer VHR imagery at 1.5 m pan, 6 m multispectral with 4 bands.

VHR sensors, especially those arranged in constellations, can offer high revisit times (e.g. daily), which support near real-time monitoring applications, including disaster mapping. However, the higher spatial resolution of VHR generally requires a smaller field of view, and thus less spatial coverage per scene (Andries *et al.* 2021). For example, WorldView-2 has a 16 km swath width. VHR imagery can offer improved opportunities for habitat mapping especially concerning finer spatial scales and achieving greater thematic detail (Andries *et al.* 2021; Nagendra *et al.* 2013). As already noted, the primary disadvantage relates to cost, especially where country-wide coverage at regular repeat frequency is required. Furthermore, VHR data in some cases is only accessible through download portals, which is not convenient for repeated access or automated processing workflows.

Hyperspectral imagery

In contrast to multispectral sensors, hyperspectral sensors typically employ hundreds of very narrow spectral bands (10 to 20 nm), and typically offer coverage further into the shortwave infrared portion of the spectrum. In the case of habitat mapping and condition assessment, hyperspectral imaging is particularly attractive, as the improved spectral resolution can enable enhanced discrimination between the specific traits of different vegetation types and habitat classes (Lausch *et al.* 2016).

While hyperspectral imaging has existed for several decades, its use has been mainly limited to airborne platforms, and research applications. Although the first spaceborne sensor, Hyperion (NASA), was launched in 2001, satellite-based sensors have only recently started to gain more traction, with sensors including PRISMA and EnMAP launched in recent years. However, there is generally a trade-off between spectral resolution and spatial resolution, with most of these platforms acquiring data at 30 m. While this may facilitate broad scale land cover analysis, it is likely to be too coarse for detailed habitat mapping. The high dimensionality of hyperspectral imagery has presented a major computational challenge in leveraging its potential. With more recent developments in cloud computing, alongside machine learning and deep learning methods, this aspect should now be more manageable. Indeed, RF and SVMs have been demonstrated as being well suited to supporting and leveraging the explanatory power of hyperspectral datasets (Maxwell *et al.* 2018). Jarocińska *et al.* (2022) undertook a study comparing airborne hyperspectral imagery (HySpex) and simulated Sentinel-2 multispectral imagery for discriminating Natura 2000 habitats in Poland. Their findings note the importance of reducing the dimensionality of the data (using linear discriminant analysis) to identify the most relevant spectral bands for identifying different vegetation types, with certain VNIR and SWIR bands identified as especially valuable for their application (Jarocińska *et al.* 2022). Planned future missions, such as ESA's Copernicus Hyperspectral Imaging Mission for the Environment (Chime) (planned for launch in 2028), will likely stimulate continued development of hyperspectral analyses for habitat assessment and related applications.

3.7.3 Synthetic aperture radar

Synthetic aperture radar (SAR) is growing in use for biodiversity assessment, particularly as a complementary source to optical imagery (Lausch *et al.* 2016). This is especially evident since the advent of Sentinel-1, where SAR has become freely available at point of access, with high revisit frequency (similar to Sentinel-2). SAR offers two primary advantages which can add value when combined with optical imagery for habitat mapping (Lausch *et al.* 2016; Nagendra *et al.* 2013):

- As an active microwave technique, SAR is unaffected by cloud, which means it can supply information at high temporal frequency.
- SAR can provide information on vegetation 3D structure and is valuable for separating land cover types which may be spectrally similar but structurally different.

Alongside Sentinel-1, commercial SAR platforms and constellations such as TerraSAR-X, COSMO-SkyMed, ICEYE, and NovaSAR, can capture radar data at higher spatial and temporal resolutions, with many of these systems targeted at high resolution monitoring. However, like VHR optical sensors, cost of data access can be high. Furthermore, as an active sensing technique, SAR presents a more complex data source in comparison to optical imagery. SAR systems emit a microwave signal which travels through the atmosphere and is reflected from the Earth's surface, with a portion of the signal received by the sensor. Understanding how best to process SAR data and extract meaningful information presents a challenge to the user community. SAR wavelength influences depth of penetration into surface vegetation, as illustrated in Figure 3. Longer wavelengths (e.g. L-band) are more suitable for deeper penetration and the likes of forestry applications, while C-band SAR (e.g. Sentinel-1) offers a compromise wavelength, achieving moderate vegetation penetration, which supports a wide range of applications, (NASA EarthData 2023).

The use of SAR for land cover and habitat mapping has seen increasing uptake (Lausch *et al.* 2016). Schmidt *et al.* (2018) combined Sentinel-1 SAR with Sentinel-2 multispectral imagery to assess the quality of dwarf shrub heathland in Germany, achieving accuracies of 76% and 73% at two different sites. Textural variables were derived from the SAR imagery and used to aid in separating vegetation classes, with the authors highlighting the value of SAR as complement to the optical imagery (Schmidt *et al.* 2018). Living England, Living Wales and JNCC's Northern Ireland (Living Maps) habitat mapping make use of Sentinel-1 (backscatter, coherence) within their classification approaches.

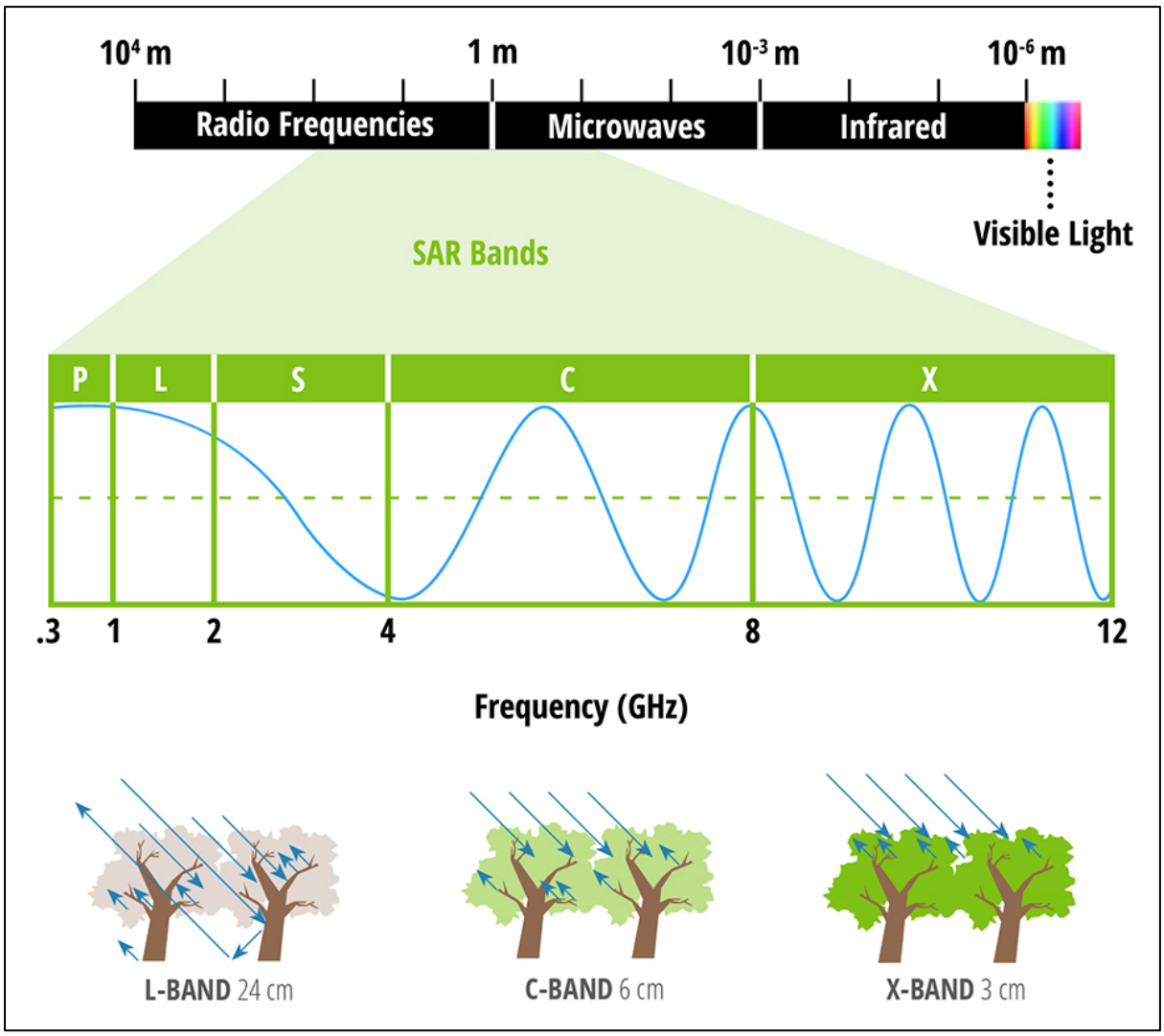

Figure 3. SAR microwave bands as part of the electromagnetic spectrum (middle and top) and highlighting typical signal penetration depth into vegetation canopy at different frequencies (bottom). Adapted from NASA EarthData (2023).

3.7.4 LiDAR

In a similar vein to SAR, LiDAR (Light Detection And Ranging) is an active EO technique which can offer complementary information to optical sensors on 3D vegetation structure (Schmidt *et al.* 2018). Spaceborne LiDAR systems are not generally optimised for habitat mapping, but airborne LiDAR is widely used for a range of environmental applications. LiDAR sensors emit high frequency laser pulses which are partly reflected from the vegetation canopy, and partly from the ground surface. By measuring the returned signals, it is possible to generate both digital terrain models (DTM) (bare surface model) and digital surface models (DSM) (top of canopy). By subtracting these, a canopy height model (CHM) can be produced. CHMs are increasingly being used in forestry, vegetation and habitat mapping, and the elevation models (e.g. DTM) can also supply topographic parameters (e.g. slope, aspect).

LiDAR is an attractive data source for habitat mapping due to the high spatial resolution of the data, which can provide additional context for habitat classification. In raw form, LiDAR is a ‘point cloud’ and point densities of 30 points per m^2 are readily achievable. However, given the huge data volumes generated, data is commonly converted to a raster DTM and DSM with 1 m or higher spatial resolution often available for users. LiDAR data has been acquired

for selected parts of Scotland on a phased basis over several years. While Scotland-wide coverage is currently not available, there are aspirations to achieve this in the near future. Acquiring country-wide LiDAR coverage for Scotland is a significant undertaking, given the size of area to be captured, and constraints posed by weather and potentially tidal conditions. Achieving full coverage is likely to take a minimum of three years. However, the value of LiDAR for habitat mapping and assessment is already being demonstrated in other UK countries. England has a country-wide LiDAR survey programme run by the Environment Agency, which implements rolling national capture ensuring continuing availability of updated coverage. Wales meanwhile has recently completed a programme of country-wide LiDAR capture. In both cases LiDAR data is becoming increasingly important to national-scale habitat mapping through the Living England and Living Wales programmes. As the update frequency for LiDAR data tends to be low (e.g. several years between repeat capture), it is generally best-suited as a baseline dataset, or for use in specific aspects of a classification, where features are likely to be stable over long periods of time.

Brownett and Mills (2017) undertook a rule based OBIA classification of sand dune habitats in England, combining LiDAR (CHM and slope) and airborne multispectral imagery. They report an overall classification accuracy of 84 %. LiDAR has also been used in combination with VHR to develop indicators of riparian forest habitat quality (Riedler & Lang 2018) and is very commonly applied in forestry applications. The Living Wales land cover and habitat maps make use of LiDAR, using a CHM to aid in refining woody vegetation classes (Planque *et al.* 2020), and the value for such purposes are discussed at length by Jongman *et al.* (2019). Living England are actively investigating greater inclusion of LiDAR in their habitat classification process.

3.8 Ancillary data

Habitats are identified by abiotic environmental influences, such as climate, geomorphology, and soils, as well as plant species composition (Agrillo *et al.* 2021; Moss 2008). Thus, it follows that such environmental factors will have a significant influence on habitat distribution. As noted in Section 3.2.3, machine learning classification approaches offer the advantage of being able to accept additional non-EO datasets which may be able to aid in habitat discrimination. The inclusion of ancillary data alongside EO data has been demonstrated as improving the overall classification accuracy (Pesaresi *et al.* 2022). Commonly used ancillary datasets include the following:

- Geology and soils.
- Climate (e.g. rainfall, temperature, humidity).
- Topography (e.g. elevation, slope, aspect, height above drainage).

Approaches such as RF allow for exploratory classification and ranking of input variables in terms of their explanatory power. Those which are identified as having little or no influence can then be removed.

Land cover and habitat mapping may also make use of other datasets to help reduce uncertainty in the classification process. This may typically include digital cartographic datasets which can be used to delineate certain well-defined classes, such as built-up areas, water, and classes at the coastal zone (e.g. Living England, Living Wales).

Other data used in the habitat classification process includes reference data used for training and validation. This aspect is discussed in Section 5.2.

3.9 Seasonality

Capturing seasonal variations related to plant phenology can be very valuable in improving discrimination between different classes and enhancing classification results (Morton & Rowland 2015; Mikula *et al.* 2021; Lucas *et al.* 2015). This typically translates to a requirement for capturing multiple datasets at different stages in the growing season and is particularly relevant considering optical imagery. Lucas *et al.* (2015) observed that for northern Europe (seasonal environment), pre- and peak-flush images should ideally be acquired in early spring and mid-summer, where these relate to the stages of lowest and highest vegetation productivity. However, due to the challenge of cloud cover, it is not always possible to capture imagery at the ideal time (Nagendra *et al.* 2013).

Another related consideration is the scale over which mapping is being undertaken, and whether a single classification model can faithfully model habitat classes over larger spatial extents. For example, in the Living England approach, England is divided into 14 regions, termed biogeographic zones. These different zones are intended to account for phenological and habitat variation, and to facilitate the production of cloud free Sentinel-2 mosaics (Kilcoyne *et al.* 2022).

Cloud, shadow or atmospheric effects can contaminate optical imagery (e.g. Sentinel-2) in an unpredictable way, as can short-term land management activities. Consequently, image mosaics are often created. These are composite images which are generated by combining the ‘best pixels’ from multiple images within a restricted window of time, allowing the effects of cloud and other contaminants to be removed. This composite image can then be used to provide a ‘seasonal mosaic’ to represent a particular period (season) in time. For instance, Living England uses four-week periods in spring and autumn to create best pixel cloud-free Sentinel-2 mosaics. Temporal filtering is also applied to Sentinel-1 backscatter imagery over these same periods to reduce radar speckle (noise) (Kilcoyne *et al.* 2022).

4 EO-based habitat mapping across the UK

4.1 Overview

While this report was commissioned by NatureScot and Scottish Government, it is relevant to look more widely at how other UK countries are addressing EO-based habitat mapping and what progress has been made to-date.

4.2 Scotland

NatureScot has been developing and publishing a Habitat Map of Scotland (HabMoS). This process has involved translating existing habitat survey data from sources such as NVC, to EUNIS classes, to publish a map offering improved country-wide coverage. However, there remains a significant gap in upland areas, which are not well-covered by existing surveys.

This led to work around 'filling' this gap and included the Upland Mapping Pilot (Scobie 2018) mentioned in Section 3.7.1. This explores stereo colour near-infrared aerial photography for mapping of uplands, focussing on two test sites. Upland areas in Scotland present a particular challenge in terms of habitat mapping, due to their extensive and remote nature. This project augmented the conventional API approach by using OBIA and undertaking image segmentation as a basis for delineating habitat boundaries. Manual stereo interpretation of the segmented images was then employed to decide whether to accept, merge or spilt individual polygons, with a view to classifying habitats at EUNIS level 3 or at greater detail (level 4 or 5). This approach involved relatively high levels of field survey to clarify uncertainties in the API interpretation, and for purposes of validation and accuracy assessment. The study drew positive conclusions regarding the efficiency of OBIA. While the approach was deemed cost-effective in comparison to fully-field based mapping, costs were estimated at £1.8 million and taking six years to complete the project with five dedicated mapping officers (Scobie 2018). This is likely significantly more costly than the kind of automated or semi-automated EO-based approaches discussed in Section 3.2.

More recent developments include the release of NatureScot's Scotland-wide land cover maps for 2019 and 2020, mapped to EUNIS Level 2. This 'SLAM-MAP', was produced by an external contractor, and classifies 22 land cover classes at a spatial resolution of 20m, delivering a pixel-based output and employing a machine learning method. This is intended to be a complete and repeatable map of Scotland and used data from the HabMoS for training and validation. An overall accuracy of 91 % is reported by the contractor, with NatureScot currently prioritising an independent field-based accuracy assessment, to provide additional insights.

4.3 England

The Living England project, led by Natural England, delivers England-wide habitat probability maps based on EO data. This approach is based on OBIA and applying a RF classifier. Segmentation of Sentinel-2 imagery is undertaken to derive land parcel objects, with cartographic and thematic datasets (OS OpenData, EA saltmarsh extent, RPA Crop Map of England (CROME)) incorporated to assign certain classes directly. A RF classification then uses Sentinel-1 and Sentinel-2 imagery, incorporating ancillary data on geology, topography, and climate to predict habitat classes, based on a modified version of the UKBAP Broad Habitats. Uncertainty is expressed as habitat probability and is a measure of agreement between the classified data and validation data, with the most recent maps (Phase IV 2022) achieving an overall habitat probability of 88%, noting that there is a high degree of variability across different biogeographic zones and different habitat types (Kilcoyne *et al.* 2022). Validation is undertaken through comparison against various independent datasets,

including data collected through field survey. There are plans to make greater use of LiDAR data in future versions to aid in discrimination of some classes. Natural England intend to release updated versions every two years, thus delivering a costed, operational mapping programme which will support ongoing assessment of habitat change. The production of these maps has been enabled through cloud computing approaches and the data are freely available through an Open Government License.

4.4 Wales

Welsh Government has funded the Living Wales programme, which produces land cover and habitat maps for Wales through EO approaches. The fundamental product is a 10 m land cover map based on the FAO LCCS classification system. The method is founded on EO data including Sentinel-1, Senitinel-2, and LiDAR (CHM). These are combined with a range of additional datasets, including Open Street Map, Forest Research's National Forest Inventory, and Copernicus High Resolution Layers. The data are brought together and through a range of EO and analytical methods, are converted into 'environmental descriptors' as detailed by Planque *et al.* (2020). This allows the generation of land cover maps at FAO LCCS Level 3 and forms the basis for producing habitat maps based on a translation to Phase 1-equivalent habitats. An associated change taxonomy (Lucas *et al.* 2022) has been proposed for interpreting change over time (refer to Section 6.2). The Living Wales programme also incorporates the EarthTrack app which enables field data collection. This supports the validation of land cover maps and habitat maps, as well as gathering data for training of the land cover classification model. The Living Wales system is currently hosted on a supercomputing facility and is being made available through an open data cube, facilitating bespoke user access.

4.5 Northern Ireland

JNCC have been working with Northern Ireland Environment Agency (NIEA) to produce habitat maps for Northern Ireland. This is adopting a Living Maps approach, which is the basis for the Living England mapping. Similar to the Living England approach, objects are generated through segmentation of Sentinel-2 imagery, and then classified using a RF algorithm, using variables derived from Sentinel-1 and Sentinel-2 seasonal mosaics (band values, band ratios and indices such as NDVI and EVI) as well as ancillary data comprised of geology, topographic variables (slope, aspect, elevation, etc.), climate (rainfall, humidity, temperature, frost), and proximity to features including coast, roads and water. The assigned habitat classes are derived mainly from the NVC classification system, with a total of 25 classes. Training and validation data is derived from NVC field data from surveys of areas of special scientific interest (ASSIs) and visual interpretation of imagery, aiming for at least 50 points per class. This has produced habitat maps for four distinct zones across Northern Ireland based on underlying geology, with accuracies ranging from 73% to 88% and expressed as mean probability per class. Work is ongoing to continue to improve the accuracy of the mapping, with challenges relating to relatively outdated field samples, a lack of validation data for locally rare habitats, and significant over-prediction of arable classes.

4.6 UK level

UKCEH produce annual UK-wide land cover maps at 10 m spatial resolution, with a legacy dating back to the Land Cover Map 1990 (LCM1990). Maps are available for 1990, 2000, 2007, 2015, 2017, 2018, 2019, 2020 and 2021. Earlier maps have been revised to conform to the current approach in terms of spatial coverage, spatial framework, and habitat classification, with an emphasis on enabling land cover change detection (Morton & Rowland 2015). The LCM are based on modified versions of the UK BAP Broad Habitat classes, with 21 classes in total, and are similar in level of detail to the Living England habitat maps. They

are the only UK level products offering a consistent classification for the whole country. Similar to other products described above, the LCM implements a RF classification. However, it differs from some in that a pixel-based classification is undertaken rather than an object-based classification. Another key difference is the use of generalised digital cartography (OS MasterMap) as the spatial framework. The pixel classification is then aggregated for the land parcel framework, thereby providing mapping at parcel (object) level as well as pixel level. This provides a fixed set of land parcels which enables analysis of change over time at a land parcel resolution but does not easily represent changes at a sub-land parcel scale (Barber & Robinson, in press). The classification uses Sentinel-2 seasonal composite images and ancillary data including topographic variables, proximity to features including roads, buildings, sea, and freshwater, as well as digital mapping data to define coastal areas, woodland and urban areas. The LCM2021 has a stated overall accuracy of 82.6%, and was validated using a composite reference dataset. This is comprised of data from the GB Countryside Survey, National Forest Inventory data, Rural Payment Agency data, and a set of LCM validation points generated from field data collection and manual image interpretation (Marston *et al.* 2022). The data aims to represent land cover. As such, differences between the LCM products and other products aiming to represent habitats need to be considered in terms of the input data used to train the classification algorithms, which will strongly influence the output classes. This should be considered by users when deciding what product would be most appropriate for their application.

5 Uncertainty

5.1 Overview

In any mapping project, it is necessary to understand sources of uncertainty, how these can be minimised, and ensure that they are clearly communicated to users. Ideally, an acceptable level of accuracy should be defined at an early stage, considering the purpose of the map, and how it will be used (Pesaresi *et al.* 2022). This section summarises the main sources of uncertainty related to habitat mapping.

5.2 Reference data

Reference data are essential not only for validating the accuracy of EO classification, but also for identifying classes in a supervised classification strategy or training a machine learning classifier (refer to Section 5.2.4). Reference data are typically considered to be 'perfect' and error free and are often referred to as 'ground truth', but this is rarely the case (Foody *et al.* 2016).

Acquiring reference data is usually the most expensive and time-consuming part of EO-based analysis (Morton & Rowland 2015). Field surveyed reference data are generally considered the ideal approach. While API is often used as a lower-cost alternative for generating reference data, the resultant classifications are often highly subjective between interpreters (Pesaresi *et al.* 2022; Powell *et al.* 2004). Pesaresi *et al.* (2022) suggest that UAVs could be useful in acquiring reference data for specific habitats which may be inaccessible.

The positional accuracy of reference data is important to consider. This should generally be an order of magnitude better than the accuracy of the EO and ancillary datasets, and it is best practice to use global navigation satellite systems (GNSS) positioning to ensure accurate georeferencing of field reference data. Related to this, the co-registration of the classification (EO and ancillary) datasets should also be inspected to resolve any offsets or misalignment which will introduce further uncertainty.

Furthermore, data collection in the field should take into consideration the spatial resolution of the EO data, or the size of segmented image objects in the final classification and the MMU. A 10 m Sentinel-2 pixel may contain a mixture of habitats, while a reference sample in the field will likely refer to a single habitat (Barber and Robinson, In press).

When land cover and habitat mapping is carried out over large extents (e.g. regional or national scale), it is often the case that reference data is drawn from multiple different sources, which can introduce various biases. For example, as discussed by Foody *et al.* (2016), different sampling strategies may have been used, some data may be contributed through more informal citizen science approaches, while other data may follow more formalised survey protocols. Further uncertainty may arise from when the data was collected and how up to date it is with respect to the EO data (Nagendra *et al.* 2013). These aspects all introduce uncertainties which are difficult to quantify and propagate into overall uncertainty metrics.

5.2.1 Spatial framework

In a spatial framework derived through OBIA, the quality of image segmentation will influence how faithfully the objects delineate habitat units on the ground, and therefore the quality of the resultant habitat map. The objective of the segmentation is to derive the optimal segmentation to achieve the highest possible classification accuracy (Costa *et al.*

2015). The segmentation is usually derived through trial-and-error testing of parameters and assessed through visual inspection in relation to the defined MMU. This process will be inherently influenced by the imagery and may be affected by aspects such as time of year and illumination conditions (Scobie 2018). The representation of mosaic habitats, small, fragmented habitats and habitat transitions needs to be considered in the segmentation design. For instance, the Upland Mapping Pilot for Scotland allowed multiple habitat classes to be assigned to a single polygon object. In this sense, a pixel-based approach, such as that used by UKCEH in the LCM can be considered to offer greater flexibility in preserving the finer scale details of mosaics or fragmented habitats. These considerations are also important for approaches which utilise a spatial framework based on digital cartographic data sources (e.g. UKCEH LCM) as there is always a trade-off between the ideal level of detail for mapping (MMU) and the ability to accurately represent all aspects of the landscape. Further, the ability of the cartographic entities to faithfully represent habitat units must be considered, as well as awareness of potential licensing restrictions.

5.2.2 Habitat characteristics

The nature of habitats themselves play a role in uncertainty. Classification systems are an attempt to describe and classify a very variable concept, to impose order on it and to create a simplification. Beyond the case of mosaic or transitional habitats, there are certain habitats which are challenging to classify through EO approaches. This is particularly the case for habitats which are spectrally similar to others, which demonstrates why considerable care is needed in selecting the definitions of classes that are being used. While not encompassing more recent machine learning approaches, the Crick Framework (JNCC 2015) provides useful context on how applicable different EO techniques are for identifying individual BAP Priority habitats and Annex I habitats.

As discussed in Section 3.9, incorporating imagery at higher temporal frequency can sometimes improve discrimination of spectrally similar classes. Increased spectral resolution can also improve the identification of certain classes. Jarocińska *et al.* (2022) demonstrate the value of hyperspectral imagery over multispectral Sentinel-2 in improving the classification accuracy of meadows and grassland, indicating that for some habitats, certain sensor types may add additional value in the EO classification process and reduce uncertainty.

5.2.3 Classification process

The selection of EO classifier, and how this is applied will influence classification accuracy. The advantages offered by machine learning approaches have been discussed in Section 3.2.3. Generally, these have been found to offer improved accuracies over conventional classification methods, especially in the case of complex datasets with many predictor variables (Maxwell *et al.* 2018). However, selecting and implementing a machine learning classifier is not necessarily straightforward, and most approaches require users to define certain parameters. Maxwell *et al.* (2018) provided a detailed review of machine learning classification for EO, focussing on practical considerations around their implementation. They highlight RF and SVM as performing especially well under a range of scenarios, and RF as being the most straightforward to implement overall.

There are specific considerations around the selection of reference data in the context of machine learning. Typically, 70-80% of the reference data is used for training the model, while 20-30% is reserved for internal cross-validation (e.g. Agrillo *et al.* 2021; Kilcoyne *et al.* 2022). However, there are several aspects around selecting appropriate reference data. Machine learning classifiers are known to be relatively sensitive to class imbalance, where the number of samples varies greatly between classes, and can cause under-prediction of less abundant classes (Maxwell *et al.* 2018). Capturing reference data for rare habitats,

which may often be the most interesting from a conservation perspective, requires specific consideration (Merrington *et al.* 2021). Stratified random sampling is generally recommended to avoid class imbalance (e.g. Agrillo *et al.* 2021; Schmidt *et al.* 2018). Maxwell *et al.* (2018) provided a discussion around scenarios where this may not be feasible and highlight alternative sampling options which could be considered.

In terms of uncertainty, the RF method is known to be sensitive to the size and representativeness of the training sample, as well as noise in the data (Belgiu & Drăguţ 2016). This reinforces the requirement for robust consideration of reference data, paying attention to sampling design and minimising class imbalance. Another aspect discussed by Belgiu and Drăguţ (2016) is transferability, highlighting that accuracies may decrease when the classification model is applied to a different area or site to the one which was used for training. A similar impact is reported by Schmidt *et al.* (2018) who tested the transferability of a SVM model using combined Sentinel-1 and Sentinel-2 datasets to classify dwarf shrub heathland. They reported a slight decrease in accuracy when the model was applied to the same site, but using data acquired over a slightly earlier time period. Such considerations may be especially important where a classification approach is being applied over regional or national extents (e.g. hence the separate biogeographical zones applied across Living England habitat maps). For further discussion around the practical implementation of RF, refer to Belgiu and Drăguţ (2016).

5.3 Communicating uncertainty

Most EO classification analyses rely on the error (confusion) matrix (Congalton 1991) to communicate classification accuracy. This compares the assigned (classified) data against reference data (sometimes called validation data). In the case of machine learning classifiers, where a proportion of the reference data is reserved for estimating error and variable importance (e.g. RF), it is important to perform an independent accuracy assessment using separately withheld reference data (Maxwell *et al.* 2018).

The error matrix delivers overall classification accuracy, as well as producer's accuracy, user's accuracy, commission, and omission error (Congalton 1993). The error matrix is well-established and allows relatively straightforward communication of map accuracy for end users. However, there are potential drawbacks which should be considered.

A figure of 80% is often considered as a threshold for acceptable overall accuracy (OA) in an EO classification (Pesaresi *et al.* 2022). While this may be useful in the broadest sense, OA is a summary of the combined accuracies across all classes and will likely mask variations in accuracy between individual classes. By examining classes with notably high or low accuracies, it is possible to gain enhanced insights into performance. For example, classes with lower accuracy may highlight shortcomings which could relate to an imbalance of training data for these classes, or challenges in adequately representing small or rarely occurring classes. Or indeed, it may indicate that the available data is poor at discriminating these particular classes. Furthermore, examining instances where one class is frequently misclassified as a different class may highlight class confusion, which could be due to such classes being spectrally similar and difficult to separate. Exploring the error matrix in this way can help to refine and improve the classification model. It may point towards incorporating higher spatial or spectral resolution data, or better representing seasonal variations to improve class separability. Some studies take the approach of using the accuracy assessment to drive improvements in the classification process by identifying classes where more or improved reference data is required (Merrington *et al.* 2021).

The error matrix also fails to communicate positional uncertainty. This relates to discussion in Section 5.2.1 around positional accuracy and co-registration of reference and classification data, but also concerns the spatial distribution of error. Classification errors are

not usually randomly distributed but tend to be correlated to class boundaries or heterogeneous regions which may be affected by mixed pixels (Foody 2002). Barber and Robinson (In press) present further discussion on the error matrix, and particular challenges around representing uncertainty in habitat change mapping.

The Living England habitat map is presented as a habitat probability map and includes details of predicted habitat probabilities from the RF classification. Table 2 summarises probabilities for the Living England biogeographical zones, but probability can also be communicated on a spatial basis for individual habitat units (e.g. primary predicted habitat, with associated probability; secondary predicted habitat with associated probability) (Kilcoyne *et al.* 2022). Ongoing work by JNCC in undertaking habitat mapping in Northern Ireland is exploring how per-class probability can be visualised to communicate how uncertainty varies spatially, following approaches proposed by Fisher (2010).

Table 2. Summary of overall probabilities for each biogeographical zone as calculated by the random forest models for Living England Phase IV (Kilcoyne *et al.* 2022).

Biogeographic Zone	Probability
BGZ01	86.64
BGZ02	92.34
BGZ03	78.73
BGZ04	92.26
BGZ05	85.65
BGZ06	91.40
BGZ07	92.73
BGZ08	90.04
BGZ09	92.12
BGZ10	85.21
BGZ11	90.37
BGZ12	81.82
BGZ13	89.71
BGZ14	89.71

6 Habitat condition and change

6.1 Overview

The Nature Positive 2030 report was published by the four countries of the UK in 2021 (Natural England *et al.* 2021). It laid out the need for evidence and targets to help the UK meet its nature goals that will help address the twin crises of climate change and biodiversity loss. This has been further underpinned with new international targets agreed under the Convention on Biological Diversity (CBD) at the end of 2022. As investment in restoring and improving our natural capital increases to meet these targets it will become increasingly important to gather evidence on the changing landscape. EO can play a part in this evidence base, particularly as it has the benefit of providing data across whole regions, but this is a new approach and so methods are not yet fully developed.

Using EO methods for assessing the extent of habitats has already been discussed but an equally important aspect of these assessments is understanding when habitats are in good condition. Well-performing habitats will be more resilient to external impacts and will support a healthy assemblage of species. The benefit of EO in this context is that we now have access to regularly repeated data collection giving the potential to analyse how habitats are behaving throughout the year, which in turn can provide insight into the function or management of that habitat.

Habitat change and habitat condition are distinct concepts but intrinsically linked, depending on how we analyse the data, so it is important to be clear on how we are using these terms. Within this report we are using the term habitat change to describe the process of identifying a location that has changed from one habitat class to another between two points in time, potentially between two classified data products. Habitat condition is describing the state of a certain habitat, but the process of using EO to determine condition is likely to look at how that habitat is behaving over time, which will involve reviewing how it is changing with the seasons or with management, but not changing into a different habitat class.

6.2 EO for assessing habitat change

As already stated, this will focus on identification of locations that have changed from one habitat class to another. There are different approaches in identifying change using EO, including map-to-map, image-to-image, or map-to-image. Each approach has its strengths and weaknesses depending on the questions being asked and the data available.

A common approach to understanding habitats with EO is to produce classified data products, which we generally refer to as maps, although they can be more accurately described as models. If two such models are created for the same location but at different times, then these can be compared to identify the differences, and this is a map-to-map approach. On a technical level such comparisons are simple, but considerable care needs to be taken in doing so. Both maps are generalisations of reality and may have been produced using different methods meaning that a user needs to understand how that would impact on the change analysis. In many cases maps created using different methods should not be considered comparable, as it would be difficult to determine whether changes are genuine or introduced through differences in the way that the maps have been produced. Both maps are accurate to a certain level and will contain some degree of uncertainty. Again, it is difficult to differentiate between changes associated with real change on the ground, as opposed to changes which may be due to errors in either or both maps propagating into the change analysis. Accuracies and uncertainties in the data should therefore be taken into consideration in any such change analysis. Map-to-map assessments of change are further considered in Barber and Robinson (in press).

Image-to-image change detection methods can work on a pixel or an object basis. When working at a pixel level, then it is vital to ensure that the pixels are perfectly aligned between images, so there are no issues with spatial misalignment. If objects are being used, then data from the image needs to be summarised for that object using statistical or modelling approaches. NDVI can be a good value to use as it normalises for differing illumination across the image. In some cases, users will remove pixel values from edges of objects to reduce impact of edge features on the comparison. This could add considerably to the processing but could also improve accuracy of the outputs.

Map-to-image methods utilise the knowledge from the map regarding the habitat at that location and will identify signals in later images that are different to those anticipated for that habitat type. This can offer greater flexibility than map-to-map approaches where the analysis is dependent on two comparable data products being available. When using this approach, the user is not constrained by the method used to create two data products and require them to be compatible as required when using a map-to-map approach. This option gives the user the ability to define their own analysis method that is appropriate to the image and map data available.

Lucas *et al.* (2022) presented a global change taxonomy, based around the driver-pressure-state-impact-response framework. This attempts to assign meaning to changes determined through EO-derived land cover maps, by use of a change glossary and identifying related impact (pressure) categories. It therefore provides a framework to aid interpretation of change regardless of the method used to identify it. This approach is being developed as part of the Living Wales land cover and habitat maps initiative.

A further example is the Land Evaluation Tool being developed by JNCC which is based on assessing changes to land parcels over time. Unlike the methods described above this aims to present time series of data to users showing the variability of each parcel of land over that period. It utilises vegetation indices derived from extended time-series of Sentinel-1 and Sentinel-2 imagery and can be used to highlight anomalies which may relate to changes in management practices or as a result of other change processes (Lightfoot *et al.* 2021). By relating the time series of data to a habitat classification and by presenting the index values graphically with the satellite imagery alongside, it facilitates the user interpreting the variability over time, potentially by the application of the Lucas *et al.* (2022) change taxonomy. Recent work with NatureScot explored the potential of this as a tool for detecting and assessing habitat change at selected protected sites in Scotland (Black *et al.* 2023). The SLAM-MAP, Living England, Living Wales and UKCEH LCM approaches are all working towards, or have already started generating change maps, based on year-to-year changes in the respective broad habitat and land cover maps.

All of approaches require an understanding of the features that are being analysed as well as the data available to the analysis.

6.3 Habitat condition

Approaches to assessing habitat condition are not yet well developed, although a wide variety of approaches are being explored. Some focus on specific aspects of the condition of certain habitats such as thermal signatures of peatlands (Worrall *et al.* 2020), or the surface movement of peatlands (Bradley *et al.* 2022). Both approaches look to relate these measurements to particular condition states of the habitat. Others look to combine a series of different measurements that relate to aspects affecting the overall condition, then bring them together to make an overall assessment. An example of this, again relating to peatland is the Peatland Restoration Portal being developed by Environment Systems as a CivTech Challenge. At regional scale (whole-EU) and with an emphasis on assessing landscape fragmentation, Chetan *et al.* (2021) use EO data to analyse change in dominant land cover

class, landscape structure, and vegetation greenness to understand how the condition of Natura 2000 sites have changed over several decades. In another example, in this case focussing on riparian forests, Riedler *et al.* (2018) derived several spatio-structural indicators from EO data to develop habitat quality metrics for informing management actions. Schmidt *et al.* (2018) combine Sentinel-1 and Sentinel-2 data through a SVM approach to translate rule-based field guidelines for assessing dwarf shrub heathland quality into EO quality layers. These examples demonstrate that while a number of initiatives and studies have sought to develop EO-based approaches to assess habitat condition, these tend to be specific to particular sites or habitat types, or in measuring particular aspects of condition (e.g. habitat fragmentation). Nevertheless, these examples offer a potential basis to explore the development of condition metrics which may be more generally applicable.

One such approach, which can be considered more generic, is to compile data from EO on how a habitat is changing over time, throughout the seasonal cycle. So long as the habitat type is known these signals can be assessed to see if they differ from the signal that would be expected for that habitat. This has the benefit of being able to work over whole regions and for any habitat, so long as EO data is regularly available (e.g. considering limitations as mentioned in Section 3.7.2 around cloud cover for optical imagery). The challenge lies in understanding the implications of any such changes and translating into some form of meaningful information.

6.4 Summary

Mapping the extent of habitats with EO is far more developed than assessing change or condition, which are more challenging tasks. Despite this, progress is being made towards assessing habitat change at broader scales, for instance in the case of the SLAM-MAP, Living England, Living Wales and UKCEH LCM initiatives. However, care is required in the choice of approach and understanding potential change artefacts and the propagation of uncertainty associated with change between individual map products. Assessing habitat condition is a more complex task using any approach. The key to deciding on appropriate methods is to know the questions that are being asked and the level of detail, both spatially and thematically, that would be required by the user. Once those have been adequately defined then advice can be sought or research carried out to decide whether EO approaches can meet those requirements entirely or could themselves highlight locations, habitat types or approaches, with additional information required from other sources.

7 Discussion

7.1 Developing 30x30 – current priorities

Section 1 introduced 30x30 and the process being undertaken by NatureScot to shape strategy around delivery. The outcome of initial consultations through a co-design process suggests a movement away from feature-based monitoring. There is an emphasis on transitioning to a protected or conserved area site network which delivers heterogeneity or complexity at landscape scale and merges into the surrounding landscape, rather than creating 'hostile' boundaries. The importance of effective and appropriate monitoring is recognised, using an ecosystem health approach that is focussed on management and reduction of landscape pressures to improve condition. NatureScot has summarised outcomes from the co-design process. Understanding the spatial extent of different habitats and how these are changing over time is likely to be critical to delivering 30x30.

7.2 Opportunities for EO

Drawing together the current status of EO for habitat monitoring, the following discussion highlights some of the main opportunities and considerations around integrating EO into 30x30 delivery. The first stage is around identifying the key requirements for monitoring progress towards 2030 targets. This will support understanding of how and where EO can contribute.

NatureScot has commissioned the SLAM-MAP product, which provides Scotland-wide coverage at EUNIS Level 2 and is also delivering change information. This offers the advantage of national-scale coverage, and work is ongoing to improve understanding of uncertainties, incorporating ground reference data. Considering assessment of change, as discussed in Section 6, there are complexities around comparing map products which change over time. However, the pixel-based nature of SLAM-MAP may offer advantages in avoiding the type of object-related artefacts which can arise through OBIA change assessment (Barber & Robinson, in press). The contribution of SLAM-MAP towards 30x30 should become clearer as priorities evolve, but it is likely to be of value in delivering broad-scale summaries and evaluating areas of interest for further assessment at higher levels of detail.

Stepping up in scale, it is more cost-effective and feasible to achieve EUNIS Level 3 habitat mapping over limited spatial extents rather than at national scale. This may relate to landscape scale analysis of networks of protected or conserved areas, individual sites, or specific types of habitats. Multispectral VHR imagery in combination with other datasets offers good opportunities in this regard for supporting more detailed mapping of smaller or fragmented habitat units. The Geospatial Commission have launched a trial of VHR imagery for public sector organisations across the UK and this may present a useful opportunity to explore and understand more about how VHR data could add value in Scotland. One of the main challenges is likely to be around cloud cover, and capture of imagery at the ideal time of year. Over recent years, Welsh Government have developed a strategic approach around leveraging value from EO for a wide range of public sector applications. This includes significant ongoing investment in optical VHR imagery, with access to high frequency, VHR imagery secured for the next two years, and recent Wales-wide LiDAR capture. Engaging with the Geospatial Commission trial and exploring knowledge sharing with Welsh Government could help provide insights around the value of both VHR and LiDAR.

The value of LiDAR and SAR has been highlighted for capturing structural information on vegetation, which can help in discriminating woody vegetation classes. Furthermore, as

cloud free optical imagery can be difficult to acquire, the value of SAR (e.g. Sentinel-1) for higher temporal frequency capture should be considered as part of habitat assessment.

Delivery of 30x30 requires monitoring and assessment over time. Careful consideration should be given to the design of a monitoring strategy, including the MMU, accuracy requirements, frequency of monitoring, and features or metrics to be assessed. This will inform the spatial and temporal resolution of EO data, the ancillary datasets, the spatial framework, reference data sampling design, and the classification approach. Measuring habitat condition in response to 30x30 management presents a means of inferring the impact of those actions. Where possible, a baseline dataset should be established to identify the present status of biodiversity, and to allow comparisons to measure change over time. A high-quality baseline dataset may also support a range of subsequent ad-hoc or opportunistic assessments of specific habitats or features. LiDAR data capture at national scale may offer a valuable contribution as a baseline dataset (alongside other EO and ancillary datasets), where a good quality DEM can support subsequent topographic analysis (e.g. as ancillary classification variables), or generating a baseline CHM for use in classification.

While habitat maps are valuable for a wide range of purposes, thinking beyond these to other metrics or indicators of habitat quality will likely be necessary in measuring the success of 30x30. This may include assessing habitat connectivity or fragmentation within and between protected or conserved areas, landscape complexity, or considering indicators relevant to specific habitats, all of which are elements in which EO can play a major role.

Furthermore, in this regard, there are huge efforts globally around 30x30, as countries grapple with similar questions around how to measure progress, and how EO may be able to contribute. This includes developments at global level around the CBD Post 2020 Monitoring Framework and the development of EO-based national-level indicators for measuring aspects such as habitat quality and connectivity. Keeping abreast of such developments can further inform the delivery of 30x30. This could be achieved through engagement and collaboration with other UK countries, engaging with implementation of the Defra EO Centre of Excellence Roadmap, and utilising opportunities such as the CivTech programme to explore and develop solutions.

7.3 Challenges

While this review has highlighted extensive use of EO for habitat mapping and assessment, there are a number of aspects which remain challenging.

Many of the habitat mapping studies described in the scientific literature focus on particular habitat types or are implemented across a small number of sites. While well-defined habitat types, such as forests, can be suited to mapping from EO, others present more of a challenge (Jongman *et al.* 2019). Habitats which cover small spatial extents or are less well-defined or fragmented, present particular difficulties. Other challenges include mosaic habitats (e.g. uplands) and habitats where different classes may present as spectrally similar. For these cases, higher spatial and spectral resolution data is likely required. Also, consideration should be given to the representation of mosaic habitats in any classification scheme – should the dominant habitat be assigned, or should information also be supplied on the percentage split of a number of different habitats? In the case of the latter, this may make the implementation of automated EO-based approaches more challenging.

In terms of the EO classification approach, and particularly considering machine learning classifiers, the need for appropriate and adequate training data cannot be overstated. Locally rare habitats, or those which cover very small extents can present a challenge in this regard. This aspect, in the context of training data, should be considered at an early stage in

the design of any EO-based approach. Care must also be taken around evaluating the transferability of an EO classification model if it has been developed using data for one site but is to be applied to other sites or locations.

Working with EO datasets at national scales, including processes such as segmentation and OBIA, involves huge computational resource. It is costly in terms of resources and scaling up methods over large spatial extents at high spatial resolutions, as well as storage and management of multiple, large EO datasets. High performance computing solutions, especially utilising cloud-based infrastructure makes such analyses more achievable, although care is required around implementing and documenting consistent and repeatable workflows.

As already discussed in Section 6, EO approaches for assessing habitat change or deriving condition metrics are still relatively limited. An understanding of what changes are important and how these will be assessed should be established at the outset in order to design an appropriate strategy. Further discussion around common challenges relating to EO-based change detection can be found in Barber and Robinson (in press) as highlighted in Section 3.5. Continued engagement with other countries facing similar monitoring challenges may be fruitful in providing insights into appropriate strategies, while considering the particular requirements of Scotland.

8 Conclusions and recommendations

This report has reviewed the current status of EO approaches for habitat mapping relating to delivery of 30x30, and considering habitat change and condition assessment. This highlighted that the use of EO for habitat mapping is widespread, including for the delivery of national-scale mapping initiatives (e.g. SLAM-MAP, Living England, Living Wales, UKCEH LCM). These deliver relatively broad-scale habitat or land cover maps. For example, the Scottish SLAM-MAP is based on the EUNIS classification at Level 2.

More detailed habitat mapping is achievable, with increasing emphasis on the use of VHR optical imagery, often alongside multispectral imagery, such as Sentinel-2. In addition, the integration of LiDAR (detailed vegetation structure, topography) and SAR (vegetation structure, high frequency image collection) is becoming increasingly commonplace. Due to more demanding EO data requirements and higher associated costs, detailed habitat mapping is likely to be restricted to smaller extents – for example individual sites, or a small network of protected or conserved areas.

There has been a clear shift away from conventional parametric classifiers (e.g. maximum likelihood classification) towards machine learning, with random forests the most dominant method. Machine learning can utilise many more classification datasets of differing types, including ancillary datasets (e.g. climate, geology, topography). Alongside this, many habitat mapping approaches employ OBIA for segmenting imagery into objects as a representation of habitat units on the ground. Other methods retain a more conventional pixel-based approach (e.g. SLAM-MAP, UKCEH LCM), with advantages and disadvantages to both.

A common challenge in many habitat mapping initiatives is the need for high quality, representative reference data, which is required to train the EO classification model, as well as to validate performance. Ensuring balance of reference samples across all habitat classes can be challenging in the case of small or rarely occurring habitats. Collecting field data is expensive but remains the best approach to ensure habitat maps are accurate and reliable. Planning an appropriate strategy for collecting reference data should be considered a fundamental element of the overall mapping design. This should consider how habitat units on the ground will be translated in the context of the minimum mapping unit, spatial resolution of the EO data, and the spatial framework for the mapping (e.g. based on segmented objects or existing digital cartographic datasets). Translating habitats into the selected habitat classification system (e.g. EUNIS) also requires care and consistency amongst field surveyors, as does an approach for interpreting mosaic habitats.

There are a number of methods for assessing habitat change, including map-to-map change and image-to-image change. SLAM-MAP, Living England, Living Wales and UKCEH LCM are all focussing on the delivery of year-to-year change maps, as this becomes a growing priority. Relating change to processes or management actions on the ground remains challenging, and it may be difficult to separate real change from artefacts introduced through the analysis process. Assessing habitat condition or quality is even more challenging. Fundamentally, a baseline assessment is needed for each habitat, alongside a clear understanding of what constitutes an improvement or deterioration in condition. Whilst there is an increasing literature on the use of EO for assessing habitat condition, many studies tend to focus on specific habitat types over limited extents, with methods or condition indicators which are not likely to be transferable to other habitats or locations.

Further research is recommended to understand how habitat assessments can help address priorities identified through 30x30. This may include assessing aspects such as habitat connectivity and landscape scale complexity.

Finally, computational resources and software are a major consideration around the processing, management, analysis, and delivery of habitat assessment information. The increasing availability of cloud computing resources, many of which offer embedded analytical tools, makes this more straightforward, although developing a strategy with clearly defined objectives will be important in identifying an optimum approach. This element is also likely to require resources relating to accessing the infrastructure and/or software, as well as to training of staff, or alternatively outsourcing some elements.

In conclusion, this leads to the following recommendations:

- Identify the key requirements and deliverables around 30x30 and use these to inform on the priorities for habitat mapping and assessment.
- Translate this into specifications around the scope, scale and data needs concerning EO and complementary data.
- Consider at an early stage how any EO-based approach will incorporate ground reference data to enable validation and uncertainty assessment – design from this basis upwards, considering initiatives such as EarthTrack (Wales).
- Consider incorporating LiDAR and SAR to provide additional information on vegetation structure, and to compensate for limited availability of cloud free optical imagery.
- Engage with the Geospatial Commission VHR trial to evaluate how such data may enable habitat mapping and assessment at EUNIS Level 3 and beyond.
- Evaluate requirements around computational infrastructure and software needed to undertake and deliver habitat mapping and assessment. Explore complementary initiatives (e.g. NCEO DataHub).
- Consider undertaking further research into habitat mapping outputs to determine how these can be used to support priorities identified by 30x30 (e.g. connectivity analysis, landscape complexity, softening protected or conserved area boundaries), thereby informing on the effectiveness of 30x30 actions.
- Explore the applicability of methods for assessing habitat change, including evaluating the utility of SLAM-MAP and engaging with other initiatives (e.g. JNCC's Land Evaluation Tool, Living Wales, Living England, UKCEH LCM).
- Investigate means of assessing and measuring habitat condition, including through engagement with other UK countries, and undertaking reviews and/or scoping studies to identify and evaluate promising initiatives.
- Utilise habitat assessment outputs for testing and modelling future landscape scenarios relating to potential impacts of 30x30.

References

Agrillo, E., Filipponi, F., Pezzarossa, A., Casella, L., Smiraglia, D., Orasi, A., Attorre, F. & Taramelli, A. 2021. 'Earth Observation and Biodiversity Big Data for Forest Habitat Types Classification and Mapping'. *Remote Sensing* 13 (7). https://doi.org/10.3390/rs13071231.

Andries, A., Murphy, R.J., Morse, S. & Lynch, J. 2021. 'Earth Observation for Monitoring, Reporting, and Verification within Environmental Land Management Policy'. *Sustainability (Switzerland)* 13 (16). https://doi.org/10.3390/su13169105.

Barber, R. & Robinson, P. In press. 'Scoping Object-Based Change Detection for Living England'. NCEA.

Belgiu, M. & Drăguţ, L. 2016. 'Random Forest in Remote Sensing: A Review of Applications and Future Directions'. *ISPRS Journal of Photogrammetry and Remote Sensing* 114 (April): 24–31. https://doi.org/10.1016/j.isprsjprs.2016.01.011.

Black, S., Hassall, I., Hecker, L., Mason, F., Robinson, P., Lightfoot, P., Vigano, P. & Blake, D. 2023. 'A Landscape Monitoring and Change Detection Tool for Protected Areas - towards Operationalising the Use of Sentinel Data in Scotland'. JNCC Report 730. Peterborough: JNCC. https://hub.jncc.gov.uk/assets/9d119d39-111f-4fcf-966a-b8e98b0e6b6f.

Blaschke, T. 2010. 'Object Based Image Analysis for Remote Sensing'. *ISPRS Journal of Photogrammetry and Remote Sensing* 65 (1): 2–16. https://doi.org/10.1016/j.isprsjprs.2009.06.004.

Blaschke, T. & Strobl, J. 2001. 'What's Wrong with Pixels? Some Recent Developments Interfacing Remote Sensing and GIS'. *Zeitschrift Fur Geoinformationssysteme* 14 (6): 12–17.

Bradley, A.V., Andersen, R., Marshall, C., Sowter, A. & Large, D.J. 2022. 'Identification of Typical Ecohydrological Behaviours Using InSAR Allows Landscape-Scale Mapping of Peatland Condition'. *Earth Surface Dynamics* 10 (2): 261–77. https://doi.org/10.5194/esurf-10-261-2022.

Breiman, L. 2001. 'Random Forests'. *Machine Learning* 45 (1): 5–32. https://doi.org/10.1023/A:1010933404324.

Brownett, J. & Mills, R. 2017. 'The Development and Application of Remote Sensing to Monitor Sand Dune Habitats'. *Journal of Coastal Conservation* 21 (5): 643–56. https://doi.org/10.1007/s11852-017-0504-x.

Bunce, R.G.H., Metzger, M.J., Jongman, R.H.G., Brandt, J., de Blust, G., Elena-Rossello, R., Groom, G.B., *et al*. 2008. 'A Standardized Procedure for Surveillance and Monitoring European Habitats and Provision of Spatial Data'. *Landscape Ecology* 23 (1): 11–25. https://doi.org/10.1007/s10980-007-9173-8.

Bunce, R.G.H., Bogers, M.M.B., Evans, D., Halada, L., Jongman, R.H.G., Mucher, C.A., Bauch, B., de Blust, G., Parr, T.W. & Olsvig-Whittaker, L. 2013. 'The Significance of Habitats as Indicators of Biodiversity and Their Links to Species'. *Ecological Indicators* 33 (October): 19–25. https://doi.org/10.1016/j.ecolind.2012.07.014.

Congalton, R.G. 1991. 'A Review of Assessing the Accuracy of Classification of Remotely Sensed Data'. *Remote Sensing of Environment* 37 (1): 35–46.

Congalton. R.G., 1993. 'A Practical Look at the Sources of Confusion in Error Matrix Generation'. *Photogrammetric Engineering & Remote Sensing* 59 (5): 641–44.

Cole, B., De la Barreda, B., Hamer, A., Codd, T., Payne, M., Chan, L., Smith, G. & Balzter, H. 2021. Corine land cover changes between 2012 and 2018 for the UK, Isle of Man, Jersey and Guernsey. NERC EDS Environmental Information Data Centre. https://doi.org/10.5285/027b6432-f44d-41ab-b349-89fae673d5f5

Costa, H., Foody, G.M. & Boyd, D.S. 2015. 'Integrating User Needs on Misclassification Error Sensitivity into Image Segmentation Quality Assessment'. *Photogrammetric Engineering & Remote Sensing* 81 (6): 451–59. https://doi.org/10.14358/PERS.81.6.451.

Dinerstein, E., Vynne, C., Sala, E., Joshi, A.R., Fernando, S., Lovejoy, T.E., Mayorga, J., *et al.* 2019. 'A Global Deal For Nature: Guiding Principles, Milestones, and Targets'. *Science Advances* 5 (4): eaaw2869. https://doi.org/10.1126/sciadv.aaw2869.

Fisher, P.F. 2010. 'Remote Sensing of Land Cover Classes as Type 2 Fuzzy Sets'. Remote Sensing of Environment 114 (2): 309–21. https://doi.org/10.1016/j.rse.2009.09.004.

Foody, G., Pal, M., Rocchini, D., Garzon-Lopez, C. & Bastin, L. 2016. 'The Sensitivity of Mapping Methods to Reference Data Quality: Training Supervised Image Classifications with Imperfect Reference Data'. *ISPRS International Journal of Geo-Information* 5 (11): 199. https://doi.org/10.3390/ijgi5110199.

Foody, G.M. 2002. 'Status of Land Cover Classification Accuracy Assessment'. *Remote Sensing of Environment* 80 (1): 185–201. https://doi.org/10.1016/S0034-4257(01)00295-4.

Hay, G.J. & Castilla, G. 2008. 'Geographic Object-Based Image Analysis (GEOBIA): A New Name for a New Discipline'. In *Object-Based Image Analysis: Spatial Concepts for Knowledge-Driven Remote Sensing Applications*, 75–89. Lecture Notes in Geoinformation and Cartography. Berlin, Heidelberg: Springer.

Hearn, S.M., Healey, J.R., McDonald, M.A., Turner, A.J., Wong, J.L.G. & Stewart, G.B. 2011. 'The Repeatability of Vegetation Classification and Mapping'. *Journal of Environmental Management* 92 (4): 1174–84. https://doi.org/10.1016/j.jenvman.2010.11.021.

Jarocińska, A., Kopeć, D., Kycko, M., Piórkowski, H. & Błońska, A. 2022. 'Hyperspectral vs. Multispectral Data: Comparison of the Spectral Differentiation Capabilities of Natura 2000 Non-Forest Habitats'. *ISPRS Journal of Photogrammetry and Remote Sensing* 184: 148–64. https://doi.org/10.1016/j.isprsjprs.2021.12.010.

JNCC. 2015. 'Making Earth Observation Work for UK Biodiversity Conservation - Phase 3'. Defra / JNCC Report.

Jones, E.G., Wong, S., Milton, A., Sclauzero, J., Whittenbury, H. & McDonnell, M.D. 2020. The Impact of Pan-Sharpening and Spectral Resolution on Vineyard Segmentation through Machine Learning'. *Remote Sensing* 12 (6): 934. https://doi.org/10.3390/rs12060934.

Jongman, R.H.G., Mücher, C.A., Bunce, R.G.H., Lang, M. & Sepp, K. 2019. 'A Review of Approaches for Automated Habitat Mapping and Their Potential Added Value for Biodiversity Monitoring Projects'. *Journal of Landscape Ecology(Czech Republic)* 12 (3): 53–69. https://doi.org/10.2478/jlecol-2019-0015.

Kilcoyne, A.M., Clement, M., Moore, C., Picton Phillipps, G.P., Keane, R., Woodget, A., Potter, S., Stefaniak, A. & Trippier, B. 2022. 'Living England: Technical User Guide'.

NERR108. Natural England. http://publications.naturalengland.org.uk/publication/4918342350798848.

Lausch, A., Bannehr, L., Beckmann, M., Boehm, C., Feilhauer, H., Hacker, J.M., Heurich, M., *et al.* 2016. 'Linking Earth Observation and Taxonomic, Structural and Functional Biodiversity: Local to Ecosystem Perspectives'. *Ecological Indicators* 70: 317–39. https://doi.org/10.1016/j.ecolind.2016.06.022.

Lightfoot, P., French, G., Hassall, I., Hecker, L., Jones, A., Trippier, B. & Robinson, P. 2021. 'Using Sentinel Data to Monitor Change in Habitats and Historic Landscape Features. Technical Report for the Caroline Herschel Framework Partnership Agreement for Copernicus User Uptake (Work Package Six)'. JNCC Report 687. Peterborough: JNCC. https://hub.jncc.gov.uk/assets/81189919-f015-43e1-88ab-3e5dfa1aeb83.

Lucas, R., Blonda, P., Bunting, P., Jones, G., Inglada, J., Arias, M., Kosmidou, V., *et al.* 2015. 'The Earth Observation Data for Habitat Monitoring (EODHaM) System'. *International Journal of Applied Earth Observation and Geoinformation* 37: 17–28. https://doi.org/10.1016/j.jag.2014.10.011.

Lucas, R., German, S., Metternicht, G., Schmidt, R.K., Owers, C.J., Prober, S.M., Richards, A.E., *et al.* 2022. 'A Globally Relevant Change Taxonomy and Evidence-based Change Framework for Land Monitoring'. *Global Change Biology* 28 (21): 6293–6317. https://doi.org/10.1111/gcb.16346.

Lucas, R., Rowlands, A., Brown, A., Keyworth, S. & Bunting, P. 2007. 'Rule-Based Classification of Multi-Temporal Satellite Imagery for Habitat and Agricultural Land Cover Mapping'. *ISPRS Journal of Photogrammetry and Remote Sensing* 62 (3): 165–85. https://doi.org/10.1016/j.isprsjprs.2007.03.003.

Marston, C., Rowland, C.S., O'Neil, A.W. & Morton, A.D. 2022. 'Land Cover Map 2021 (Land Parcels, GB)'. NERC EDS Environmental Information Data Centre. https://doi.org/10.5285/398dd41e-3c08-47f5-811f-da990007643f.

Maxwell, A.E., Warner, T.A. & Fang, F. 2018. 'Implementation of Machine-Learning Classification in Remote Sensing: An Applied Review'. *International Journal of Remote Sensing* 39 (9): 2784–2817. https://doi.org/10.1080/01431161.2018.1433343.

Merrington, A.T., Milodowski, D.T. & Williams, M. 2021. 'Optimising Remotely Sensed Land Cover Classification for Habitat Mapping in Complex Scottish Upland Landscapes'. In . Vol. 11888. https://doi.org/10.1117/12.2600869.

Mikula, K., Šibíková, M., Ambroz, M., Kollár, M., Ožvat, A.A., Urbán, J., Jarolímek, I. & Šibík, J. 2021. 'Naturasat—a Software Tool for Identification, Monitoring and Evaluation of Habitats by Remote Sensing Techniques'. *Remote Sensing* 13 (17). https://doi.org/10.3390/rs13173381.

Morton, R.D. & Rowland, C.S. 2015. 'Developing and Evaluating an Earth Observation-Enabled Ecological Time Series System'. JNCC Report No. 563. Peterborough: JNCC. https://hub.jncc.gov.uk/assets/f3d1df1a-1ddf-4d9c-99f1-2a378834e340.

Moss, D. 2008. 'EUNIS Habitat Classification - a Guide for Users'. European Topic Centre on Biological Diversity.

Nagendra, H., Lucas, R., Honrado, J.P., Jongman, R.H.G., Tarantino, C., Adamo, M. & Mairota, P. 2013. 'Remote Sensing for Conservation Monitoring: Assessing Protected Areas,

Habitat Extent, Habitat Condition, Species Diversity, and Threats'. *Ecological Indicators* 33: 45–59. https://doi.org/10.1016/j.ecolind.2012.09.014.

NASA EarthData. 2023. 'What Is Synthetic Aperture Radar?' 2023. https://www.earthdata.nasa.gov/learn/backgrounders/what-is-sar.

Natural England, JNCC, Natural Resources Wales, NatureScot & Northern Ireland Environment Agency. 2021. 'Nature Positive 2030 - Summary Report'. Peterborough: JNCC.

NatureScot. 2023. '30 by 30 Explained'. 2023. https://www.nature.scot/professional-advice/protected-areas-and-species/30-30-and-nature-networks/30-30-explained.

Owers, C.J., Lucas, R.M., Clewley, D., Planque, C., Punalekar, S., Tissott, Chua, S.M.T., Bunting, P., Mueller, N. & Metternicht, G. 2021. '*Living Earth* : Implementing National Standardised Land Cover Classification Systems for Earth Observation in Support of Sustainable Development'. *Big Earth Data* 5 (3): 368–90. https://doi.org/10.1080/20964471.2021.1948179.

Pesaresi, S., Mancini, A., Quattrini, G. & Casavecchia, S. 2022. 'Functional Analysis for Habitat Mapping in a Special Area of Conservation Using Sentinel-2 Time-Series Data'. *Remote Sensing* 14 (5). https://doi.org/10.3390/rs14051179.

Planque, C., Punalekar, S., Lucas, R., Chognard, S., Owers, C.J., Clewley, D., Bunting, P., Sykes, H. & Horton, C. 2020. 'Living Wales - Automatic and Routine Environmental Monitoring Using Multi-Source Earth Observation Data'. Proceedings Volume 11534, Earth Resources and Environmental Remote Sensing/GIS Applications XI. https://doi.org/10.1117/12.2573763.

Powell, R.L., Matzke, N., de Souza, C., Clark, M., Numata, I., Hess, L.L. & Roberts, D.A. 2004. 'Sources of Error in Accuracy Assessment of Thematic Land-Cover Maps in the Brazilian Amazon'. *Remote Sensing of Environment* 90 (2): 221–34. https://doi.org/10.1016/j.rse.2003.12.007.

Riedler, B. & Lang, S. 2018. 'A Spatially Explicit Patch Model of Habitat Quality, Integrating Spatio-Structural Indicators'. *Ecological Indicators* 94: 128–41. https://doi.org/10.1016/j.ecolind.2017.04.027.

Schmidt, J., Fassnacht, F.E., Förster, M. & Schmidtlein, S. 2018. 'Synergetic Use of Sentinel-1 and Sentinel-2 for Assessments of Heathland Conservation Status'. *Remote Sensing in Ecology and Conservation* 4 (3): 225–39. https://doi.org/10.1002/rse2.68.

Scobie, A. 2018. 'Habitat Map of Scotland Upland Mapping Pilot: Developing a Method to Map Upland Habitats Using Stereo Colour near-Infrared Aerial Imagery'. Research Report 983. Scottish Natural Heritage.

Sideris, K., Colson, D., Lightfoot, P., Heeley, L. & Robinson, P. 2020. 'Review of Image Segmentation Algorithms for Analysing Sentinel-2 Data over Large Geographical Areas'. Report No. 655. Peterborough: JNCC. https://hub.jncc.gov.uk/assets/43522d33-179e-40d1-864c-447fb6900e65.

Smith, G.M. & Morton, R.D. 2010. 'Real World Objects in GEOBIA through the Exploitation of Existing Digital Cartography and Image Segmentation'. *Photogrammetric Engineering & Remote Sensing* 76 (2): 163–71. https://doi.org/10.14358/PERS.76.2.163.

Stamp, L.D., 1931. The Land Utilization Survey of Britain'. The Geographical Journal, 78, 40-47.

Szigarski, C., Jagdhuber, T., Baur, M., Thiel, C., Parrens, M., Wigneron, J.-P., Piles, M. & Entekhabi, D. 2018. 'Analysis of the Radar Vegetation Index and Potential Improvements'. *Remote Sensing* 10 (11): 1776. https://doi.org/10.3390/rs10111776.

Tomaselli, V., Dimopoulos, P., Marangi, C., Kallimanis, A.S., Adamo, M., Tarantino, C., Panitsa, M., *et al.* 2013. 'Translating Land Cover/Land Use Classifications to Habitat Taxonomies for Landscape Monitoring: A Mediterranean Assessment'. *Landscape Ecology* 28 (5): 905–30. https://doi.org/10.1007/s10980-013-9863-3.

USGS. 2023. 'Landsat Surface Reflectance-Derived Spectral Indices'. 2023. https://www.usgs.gov/landsat-missions/landsat-surface-reflectance-derived-spectral-indices.

Wilson, U., Cutts, A., Jin, C., Lightfoot, P. & Jones, G. 2021. 'Trial of Open Data Cube for Environmental Monitoring with a Time-Series of Sentinel-1 and Sentinel-2 Analysis-Ready Data'. Peterborough: JNCC.

Worrall, F., Boothroyd, I.M., Howden, N.J.K., Burt, T.P., Kohler, T. & Gregg, R. 2020. 'Are Peatlands Cool Humid Islands in a Landscape?' *Hydrological Processes* 34 (25): 5013–25. https://doi.org/10.1002/hyp.13921.